U.S. Army Research Institute
for The Behavioral And Social Sciences

Research Report 1821

Novice Versus Expert Command Groups: Preliminary Findings And Training Implications For Future Combat Systems

Thomas J. Carnahan
Western Kentucky University
Consortium Research Fellows Program

Carl W. Lickteig, William R. Sanders, Paula J. Durlach, and James W. Lussier
U.S. Army Research Institute for the Behavioral and Social Sciences

March 2004

20040419 058

Approved for public release; distribution unlimited.

U.S. Army Research Institute
for the Behavioral and Social Sciences

A Directorate of the U.S. Army Human Resources Command

ZITA M. SIMUTIS
Director

Technical review by

May H. Throne, U.S. Army Research Institute
Brooke B. Schaab, U.S. Army Center, Fort Rucker

NOTICES

DISTRIBUTION: Primary distribution of this Research Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-PO, 5001 Eisenhower Ave., Alexandria, VA 22304-4841.

FINAL DISPOSITION: This Research Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Research Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) March 2004	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) Jul 02-Jul 03
4. TITLE AND SUBTITLE Novice Versus Expert Command Groups: Preliminary Findings and Training Implications for Future Combat Systems		5a. CONTRACT OR GRANT NUMBER
		5b. PROGRAM ELEMENT NUMBER 0602785A
6. AUTHOR(S) Thomas J. Carnahan, (Western Kentucky University), Carl W. Lickteig, William R. Sanders, Paula J. Durlach, and James W. Lussier, (U.S. Army Research Institute)		5c. PROJECT NUMBER A790
		5d. TASK NUMBER 211
		5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: DAPE-ARI-IK 2423 Morande Street Fort Knox, KY 40121-5620		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: DAPE-ARI-IK 5001 Eisenhower Avenue Alexandria, VA 22304-4841		10. MONITOR ACRONYM ARI
		11. MONITOR REPORT NUMBER Research Report 1821

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*

The U.S. Army's ongoing transformation to Future Combat Systems (FCS) requires a commensurate transformation in training, particularly in the area of command and control (C^2). This paper describes research comparing novice versus expert command group performance to identify training recommendations and implications. The research was performed by the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) in conjunction with the FCS C^2 program. Comparative results were based on objective measures of verbal and human-computer interaction (HCI) as well as subjective measures of workload, performance success, training, prototype effectiveness, and human system integration. Significant differences were identified between novice and expert command groups including lower verbalization and HCI rates by novices during enemy target identification, and higher rates of verbalization and HCI by novices in support of the destruction of enemy targets. Training examples were provided for more specific and tangible guidance to training developers and designers. Overall, the findings and training implications underscore the new training requirements for FCS, particularly embedded training, for transforming novices into experts in future command groups.

15. SUBJECT TERMS
Future Combat Systems   Command and Control   Training   Human Performance   Novice and Expert Performance

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited		Dr. Carl W. Lickteig 502-624-6928

Research Report 1821

Novice Versus Expert Command Groups: Preliminary Findings And Training Implications For Future Combat Systems

Thomas J. Carnahan
Western Kentucky University
Consortium Research Fellows Program

Carl W. Lickteig, William R. Sanders, Paula J. Durlach, and James W. Lussier
U.S. Army Research Institute for the Behavioral and Social Sciences

Armored Forces Research Unit
Barbara A. Black, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22304-4841

March 2004

Army Project Number 20262785A790

Personnel Performance and Training Technology

Approved for public release; distribution unlimited

FOREWORD

The U.S. Army's challenging transformation to Future Combat Systems (FCS) requires a complementary transformation in training. The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) performs proactive research focused on human performance in emerging FCS organizations to understand and meet FCS training requirements. This report describes a small but informative research effort that compared novice versus expert command group performance to identify FCS command and control (C^2) training requirements and implications.

Although it may not "take one to know one," our relatively novice first author from the Consortium Research Fellows Program brings fresh perspective and participant rapport to this novice-focused research effort. Overall, this report reflects ongoing work in the FCS C^2 program by ARI, especially the Future Battlefield Conditions (FBC) Team of the Armored Forces Research Unit (AFRU). The work supports work package (211) FUTURETRAIN: Techniques and Tools for Command, Control, Communications, Computer, Intelligence, Surveillance, and Reconnaissance (C4ISR) Training of Future Brigade Combat Team Commanders and Staffs and the Science and Technology Objective (STO) titled "Methods and Measures of Commander-Centric Training."

This report's comparison of novice versus expert performance draws from three of the five commander-in-the-loop experiments conducted at Fort Monmouth, New Jersey, as part of the FCS C^2 program, from October 2001 to March 2003. Comparative results were based on objective measures of verbal and human-computer interaction as well as subjective measures of workload, performance success, training, prototype effectiveness, and human system integration. Findings are regarded as preliminary given the exploratory nature of the research effort; however, the objective results indicated significant differences between novice and expert performance on verbal and human-computer interaction measures. Implications for training novice command group participants were identified based on the experimental results and the literature reviewed.

Findings from this effort were provided to the Program Manager FCS C^2 as part of ARI's ongoing efforts in support of the FCS C^2 program. Overall, the findings and training implications provide empirical guidance to training designers and developers for transforming novices into experts in future command groups.

STEPHEN L. GOLDBERG
Acting Technical Director

ACKNOWLEDGEMENTS

The first author wishes to acknowledge the support and contributions made by many during my work at the U.S. Army Research Institute's Armored Forces Research Unit. I thank Dr. Betsy Shoenfelt, professor and coordinator of the Industrial/Organizational Psychology Master's program at Western Kentucky University, for recommending me to the position of Research Fellow at ARI-AFRU. To Dr. Robert Ruskin and Ms. Julie Waller, from the Consortium of Universities of the Washington Metropolitan Area, for their support, and for making this paper possible by providing transportation and expense funds to work at Fort Monmouth, New Jersey, during the novice Summer Experiment. Also, to Major Ken Copeland, from the U.S. Army Communications-Electronics Command (CECOM), for his hospitality and cordial attitude during my visit to the Defense Advanced Research Projects Agency's (DARPA) Future Combat Systems Command and Control (FCS C^2) Summer Experiment. Also, thanks to Dr. Paula Durlach for contributions to this effort. Last, but by no means least, I heartily thank Dr. Carl Lickteig and Mr. William Sanders for their support and patience throughout my fellowship at ARI.

NOVICE VERSUS EXPERT COMMAND GROUPS: PRELIMINARY FINDINGS AND TRAINING IMPLICATIONS FOR FUTURE COMBAT SYSTEMS

EXECUTIVE SUMMARY

Research Requirement:

The U.S. Army's challenging transformation to Future Combat Systems (FCS) requires a commensurate transformation in training. Understanding, describing, quantifying, and measuring the differences between experts and novices will be vital to developing efficient and effective training. An ongoing research program called FCS C^2 exemplifies the Army's effort to transform command and control (C^2). This paper describes research conducted by the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) in support of the FCS C^2 program to identify training requirements and implications for small command groups.

Procedure:

From October 2001 to March 2003, the FCS C^2 program conducted a series of five commander-in-the-loop experiments at Fort Monmouth, NJ. As a key member of the Human Performance Team for FCS C^2, ARI observed and collected human performance data across experiments, including a subset of the experiments selected for a comparison of novice versus expert participants. Novice command group participants included four Army Cadets from an experiment called the Summer Experiment. Expert participants primarily included four Active Duty lieutenant colonels from Experiments 2 and 3. Comparisons were based on novice versus expert objective measures of verbal and human-computer interaction and subjective measures of workload, performance success, training, prototype effectiveness, and human system integration.

Findings:

All findings are regarded as preliminary given the research limitations, particularly the exploratory nature of the FCS C^2 program and the small sample size. Significant differences were found between novice and expert group verbal and human-computer interactions. Among those that appear to reflect true novice-expert differences are:

- Novices spent more time in silence, less time collaborating.
- Novices talked more about firing, less about seeing.
- Novices talked more about own troops, less about enemy.
- Novices talked more about enemy location, less about enemy identification and disposition.
- Novices performed fewer computer interactions to recognize and identify targets.
- Novices performed more computer interactions to assess battle damage.

Utilization of Findings:

Findings from this effort were provided to the Program Manager (PM) FCS C^2 as part of ARI's ongoing efforts in support of the FCS C^2 program. The results provide an emerging empirical database to help identify FCS command group tasks and training requirements. In particular the novice versus expert comparisons support the early identification of possible skill training requirements associated with more representative, less senior, future leaders employing FCS C2 systems. The stabilized team of highly qualified senior officers who participated in the FCS C2 program was an essential element in advancing the state of the art for FCS C^2 capabilities development. However, for successful FCS C2 systems development, particularly training system development, it was equally necessary to employ a command group composed of lesser-skilled participants more representative of the target audience personnel to ensure the early identification of performance-based system training requirements.

NOVICE VERSUS EXPERT COMMAND GROUPS: PRELIMINARY FINDINGS AND TRAINING IMPLICATIONS FOR FUTURE COMBAT SYSTEMS

CONTENTS

List of Tables

List of Figures

Page

NOVICE VERSUS EXPERT COMMAND GROUPS: PRELIMINARY FINDINGS AND TRAINING IMPLICATIONS FOR FUTURE COMBAT SYSTEMS

Introduction

As the Army works to develop the Future Combat Systems (FCS)[1] operational and training environments for command and control (C^2), researchers must identify the contribution of the expert team to observe system performance. Development efforts will not be successful if an FCS C^2 system that can only be operated by experts is designed. Early system development must identify how expert and novice personnel would perform with evolving FCS C^2 configurations. Investigating the differences and similarities between novice and experts is essential in early design of a system; even the most technologically advanced design will fail if operator requirements exceed the aptitude, expertise, and training support available. Furthermore, identifying and understanding key differences in performance between experts and novices will be essential to constructing efficient and effective training. This introductory section provides information on the FCS, the Future Combat Systems Command and Control (FCS C^2) program, and on performance and training issues pertaining to novices and experts.

The Future Combat Systems Command and Control Research Program

The Army's continuous transformation to FCS will require an unprecedented alliance of humans and machines. Development of FCS is designed to incorporate ongoing technological advances to aid human decision making and create a network-centric and progressively autonomous force. A pivotal challenge for the FCS is the requirement that a relatively small command group must be able to command and control a complex mix of manned and autonomous systems. Command and control within the U.S. Army is quickly transforming to meet FCS requirements. The FCS concept requires Commanders and Soldiers to work with robotic systems and intelligent agents to accomplish the basic C^2 functions of: Plan, See, Move, and Strike. The FCS C^2 program has the challenge of creating a system that actually improves command and control without unnecessary complexity.

The FCS C^2 experimental program was conducted by the Defense Advanced Research Projects Agency (DARPA) and the U.S. Army Communications-Electronics Command (CECOM) Research and Development Center (RDEC) at Fort Monmouth, NJ. The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) supports the FCS C^2 program as a member of the Human Performance Team. The stated purpose of the FCS C^2 program is to test the hypothesis that digitization of the current battlefield operating systems enables a new approach to command and control:

> If digitization of current battlefield operating systems can substantially enhance command and control by providing better, more accurate, and timely battlefield data to today's command and staff for decision making; then a "new" approach to Battle Command and Control, implemented in the form of synthesized/analyzed information presented to the future Unit Cell Commander, will enable him to

---

[1] A list of acronyms is provided in Appendix A.

leverage opportunities by focusing on fewer unknowns, clearly visualizing current and future end states, and dictating the tempo within a variety of environments, while being supported by a significantly reduced staff (Pronti, Molnar, & Wilson, 2002, pp. ES-3).

Figure 1. Surrogate command and control (C^2) vehicle and expert participants.

The Unit Cell was designed as a notional small combined arms unit within the proposed force structure of FCS. The FCS C^2 program developed a mockup C^2 vehicle, displayed in Figure 1, to house the Unit Cell Commander, three battlestaff managers, a driver, and a gunner. The Commander and the battlestaff managers composed the command group of the Unit Cell. The battlestaff managers supported the Unit Cell mission by performing Plan, See, Move, and Strike activities. Each member of the command group had access to a workstation with two displays that allowed them to command and control Unit Cell assets, predominantly robotic sensor and weapon platforms. Figure 2 depicts the manned and robotic elements of the Unit Cell configuration replicated in the FCS C^2 program. The Unit Cell assets include the surrogate C^2 vehicle (C^2 Veh) occupied by command group personnel, unmanned aerial vehicles (UAVs), unmanned rotary wing observation platforms (A-160s), line-of-sight (LOS) weapons, non-line-of-sight (NLOS) and beyond line-of-sight (BLOS) weapons, the Future Warriors (FW), and their vehicles.

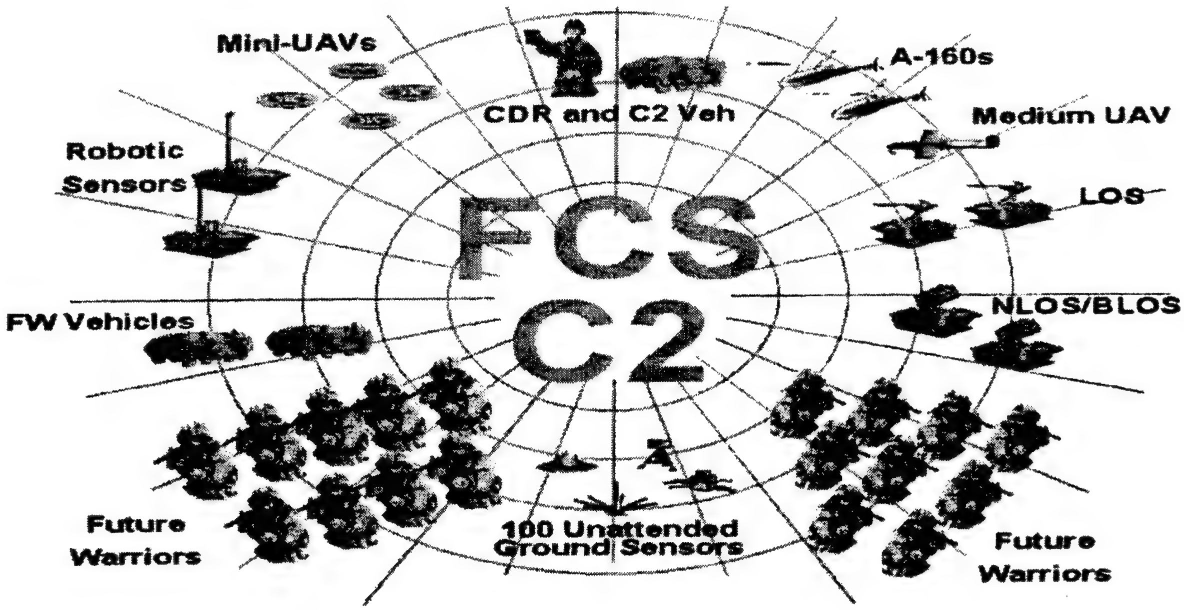

Figure 2. Organization of manned and robotic elements of a Unit Cell within the future combat systems command and control (FCS C^2) program.

Between December 2001 and March 2003, the FCS C^2 research program conducted four experiments using a highly trained and experienced team of expert U.S. Army lieutenant colonels (LTCs) as command group personnel. One of ARI's goals in this effort was to identify issues associated with effective and efficient training of command group personnel. The FCS C^2 research program furthered the training focus by conducting the Summer Experiment between Experiments 2 and 3, in which the expert command group LTCs were replaced with a novice team of Army Reserve Officer Training Candidate (ROTC) students, and cadets from the U.S. Army Military Academy (West Point).

Novice and Expert Performance

Novices are often defined as beginners or inexperienced performers in a domain, occupation, or skill being studied, while experts are defined as highly experienced performers in the domain, occupation, or skill. Researchers agree that obvious differences exist between novices and experts. Foremost and quite obviously, differences in proficiency of performance have been found in many studies (Borko & Livingston, 1989; Fisk, Kirlik, Walker, & Hodge, 1992; Kirlik, Fisk, Walker, & Rothrock, 1998). It should be noted, however, that the relationship between performance levels and amount of experience is only reliably strong when considering the novice phase of skill development. In many fields, novices rise rapidly in skill and may, after a few days, a few months, or a few years, depending upon the nature of the domain, surpass very experienced 'old hands.' For example, in chess where there is a reliable performance-based measurement system, most amateur tournament players do not ever rise above the average United States Chess Federation rating of C player despite a lifetime of experience, while many higher-rated players (A, B, Expert, and Master ratings) can be very young and comparatively less experienced. Therefore, alternative definitions of expertise based on performance rather than

experience have merit. For example, one such definition (Kramer, 1999) defines experts as those who have mastered the perceptual, motor, cognitive, and interpersonal skills necessary for performance with few to no errors.

Expertise is domain-specific. Being an expert in one domain, occupation, or skill, does not automatically transpose to other domains, occupations, or skills. However, the nature of the differences found between experts and novices are generally "consistent across a number of independent domains" (Wiggins & Henley, 1997, p. 366). Borko and Livingston (1989) found that information processing through the use of schemas by expert teachers led to more elaborate, complex, and better interconnected performance and recollection of information. Hershey, Walsh, Read, and Chulef (1990) found that accounting experts performing financial decisions for clients showed more forward reasoning strategies, in which they defined a suitable starting point and developed a strategic approach to reach the desired objective. In contrast, novices showed difficulty in creating a starting point, and, at best, created haphazard approaches to solving the accounting problem. Hershey et al. proposed that the forward reasoning used by the financial experts was indicative of a schema-based structure for organizing information. Wiggins and Henley (1997) showed that schematic organization can result in quicker and simpler examination of information and problem resolution.

Decision making is one dimension of human and command group performance that appears to vary substantially based on expertise. McKinney (1993) performed a study on crisis decision making in which he found that experts have an advantage over novices in stressful environments. The advantage of experts in decision making over novices remains even when limitations, such as time constraints, are introduced. Expert flight instructors rely mostly on schemas to aid decision making, therefore, their overall performance during crisis situations remains at a high level. This finding was mainly attributed to their ability to rely on schemas and quick adaptation to a changing and increasingly stressful environment. Wiggins and Henley (1997) studied novice and expert pre-flight decision making under stressful situations and discovered that experts tend to alter strategies in response to a change in the environment while maintaining a high level of overall performance. Successful mid-decision alteration of strategies is apparently accomplished by recourse to the vast informational schemas that experts have.

Cognitive Strategies

Thorndike and Woodworth (1901) attempted to identify a simple form of acquired knowledge. The researchers first tested participants on a mental cognitive ability, and then trained the participants on a second similar or related cognitive ability. For example, some participants in Thorndike and Woodworth's studies received training on estimating the size of a drawn square in centimeters. A similar cognitive task asked participants to estimate the size of another set of similarly drawn squares, and a related task asked participants to estimate the size of rectangles of various sizes. Thorndike and Woodworth discovered that training even similar abilities or functions had little, no, or even detrimental effects on the targeted cognitive ability. However, the researchers did notice that new strategies for making judgments in different and possibly more efficient ways were formed through training. Such early research set the stage for much of the more modern literature reviewed below on analogical reasoning, heuristics, deliberate practice, part-task training, and structured training.

Analogical reasoning. Analogical reasoning involves using past information to discover the solution to a novel situation by looking for commonalities between the previous and novel situations or, more simply, recognizing an underlying pattern between two events. The initial stage of analogical reasoning occurs when a current problem or situation primes an earlier problem or situation. The solution to the earlier problem can be adapted and expanded to create a solution for the new problem. In order for novices to be successful in this stage, they must have experiences by which to make comparisons, and they must be able to recognize the similarities between the situations.

Vosniadou and Ortony (1989) state that consideration of analogical reasoning has grown due to "the realization that human reasoning does not always operate on the basis of content-free general inference rules, but, rather is often tied to particular bodies of knowledge and is greatly influenced by the content in which it occurs" (p. 1). Hornik and Ruf (1997) claimed that the use of analogical thinking is the main difference in the cognitive processes of novices and experts.

Heuristics. Heuristics are simplified rules and mental cues that foster swift decision making. Kirlik, Walker, Fisk, and Nagel (1996) investigated the differences between novices and experts in dynamic decision making. In complex or dynamic situations, Kirlik et al. (1996) suggest that training should be aimed at teaching novices to establish routines, heuristics, and short cuts. These simplified mental cues allow the novices to cope with the demands of stressful and complex environments. Experts are able to perform better during stressful situations due to their increased use of heuristics which provide a general approach or solution to the problem, even though a variety of more specific situational constraints may exist. Experts are able to adapt heuristics and solutions from past situations through analogical reasoning to quickly solve problems in new situations.

Training Strategies

Deliberate practice. Although the benefits of ingrained habits have long been appreciated, training programs are rarely structured to fully leverage these benefits. One of many training fallacies cited by Schneider (1985), for example, is: "Practice makes perfect." A poorly structured regimen of practice, practice, practice often results in little or no improvement in learning or performance. In contrast, "deliberate practice" methods stress that training exercises should explicitly focus on relatively specific objectives and ensure repeated practice opportunities with feedback. One example of deliberate practice is the training developed by Kirlik et al. (1998) on tactical decision making with a command and control system for the U.S. Navy. The authors found that practice is necessary for proficient performance, and that novices tend to decrease in proficiency of performance during situations of high workload or demanding situations due to a lack of operational experience or practice. Another example of deliberate practice is the training provided to novice, and particularly, expert command groups for FCS C^2, as described in the Method section of this report.

Part-Task training. Part-task training involves breaking up a complicated task, a task that requires many steps for successful completion, into individual components that are trained separately. The practice and training technique performed by Kirlik et al. (1998) involved the adoption of part-task training. Two common ways for accomplishing part-task training are

segmentation and simplification. "*Segmentation* involves breaking up a task into temporal series of stages and designing part-task training around these individual stages. *Simplification* involves making the task initially easier to perform in some manner and designing part-task training for the simplified task prior to having the trainee move to the full version of the task (e.g., training wheels)" (Kirlik et al. 1998, pp. 94-95). Segmentation focuses on creating temporal step-by-step processes towards completion of the whole task. Simplification focuses on eliminating elements of the whole task that could interfere with or confuse learning and performance.

Structured training. A more structured training approach has been found to sustain proficiency of performance and improve knowledge, skills, and abilities (Campbell, Quinkert, & Burnside, 2000). Structured training is characterized by an incremental series of technical and tactical exercises. The exercises are often situated in a simulated environment to provide task realism and performance feedback. The structured sequence of training strengthens the focus on critical tasks and increases learning by building on prior learning. Campbell et al. (2000) described four different phases for creating structured training scenarios: (1) make initial decisions on the audience, assumptions and expectations for the training, (2) designate specific training objectives for the simulation, (3) design tactical scenarios to match predefined objectives, and (4) prepare training materials to supplement the tactical scenario.

Performance Feedback

The work of Kirlik, Fisk, Walker, and Rothrock (1998) relates directly to more advanced C^2 operational environments and principles important for C^2 training. They discussed four factors of performance feedback, and how each factor affected the performance of trainees on a Navy C^2 Combat Information Center (CIC). Not only was performance affected by feedback, but the authors found that training effectiveness and efficiency could be improved by enhancing feedback. Performance was affected in the following four ways:

Timeliness. Immediate performance feedback aids trainees in understanding the outcomes of certain actions, and the specific actions that led to the performance outcome. This will enable trainees to quickly and easily create and identify patterns of successful and failed performance. Kirlik et al. (1998) suggests that automated and immediate feedback be provided by the C^2 system to catch minor technical and tactical errors before they manifest into critical errors. Back briefing may not be able to identify the behaviors that preceded the critical error.

Standardization. Human trainers often have their own idiosyncratic methods for performing tasks within a given system. Kirlik et al. (1998) suggests that a strict standardization needs to be implemented to allow clear performance evaluation of the novice trainees. Trainers need to systematically tailor their feedback to trainees' basic tactical and technical skills before moving to more strategic or higher-order evaluations to increase understanding of performance information.

Diagnostic precision. Kirlik et al. (1998) suggests that trainers must be general in the feedback they provide to novices to increase performance and comprehension. Reporting to a novice a too precise performance problem, such as "you were too late in identifying the enemy platform in the Northern quadrant," may mislead trainees from the more general error made.

The overly precise and superficial feedback may not give the novice the information needed to correct the underlying problem in future training exercises. For example, a command group trainee may enter the next exercise focused on the requirement to identify all enemy targets in the Northern quadrant, at the expense of the more general requirement to maintain overall situational awareness of the battlefield.

Presentation. Kirlik et al. (1998) found that trainees on a Navy CIC system regarded most forms of auditory trainer feedback during an exercise to be distracting and unwanted. Although the information provided by the trainers was viewed as valuable to the trainees, the interruption placed more workload on already challenging duties. Kirlik et al. (1998) state that the interruptions "compromise the potential benefits that could result from expert feedback" (p. 105). Careful attention to feedback mode is required to improve versus deter novice command group performance. Literature on training novices towards the level of expert mastery was also reviewed. Training has long been an important topic among both academic and military researchers. Before successful training methods can increase learner knowledge, however, the body of knowledge to be trained must be identified and organized.

Tactical and Technical Requirements for C^2 of Future Forces

Members of Future Force command groups will need to possess an extensive mix of tactical and technical skills to achieve competent performance. Below, a list of tactical and technical knowledge and skill requirements is offered to help indicate the wide range of skills that may be required in future command groups.

Exemplar Tactical Knowledge and Skill Requirements for Command Groups:

- Knowledge of Army Brigade/Battalion staff officer duties and responsibilities.
- Knowledge of Mission, Enemy, Troops, Terrain, Time, Civilians (METT-TC) requirements.
- Knowledge of the implications of each METT-TC requirement.
- Ability to correctly order operations for all courses of action.
- Knowledge of the *why* behind the *how* in execution of system tools and features.
- Ability to visualize the battlefield from current state to desired end state.

Exemplar Technical Knowledge and Skill Requirements for Command Groups:

- Knowledge of commercial operating systems (e.g., Windows® and Linux®).
- Knowledge of alternative interface manipulation techniques to access FCS C^2 features.
- Knowledge of FCS C^2 features location and function.
- Knowledge of symbols and icons displayed by the system.
- Ability to send and receive information using the system.
- Ability to operate the communications system.

The goal of training – the achievement of expert performance levels – requires an approach that integrates the development of tactical and technical skills. Technical skills specific to a command and control system must be understood in relation to the overall tactical requirements. For example, realizing how a technical feature should be used in a given tactical

situation is required for an expert-level appreciation of the feature. Tactical skills also must be grounded in detailed technical knowledge of specific system capabilities. An accurate and detailed technical knowledge of the systems on the battlefield, their speed, power, and other inherent capabilities, is essential in order to acquire both tactical calculation skills and tactical judgment skills.

Method

This section describes the methods and measures used to assess and compare novice and expert performance. A more detailed and complete compilation of experimental methods is provided by the ARI report titled Future Combat Systems Command and Control (FCS C^2) Human Functions Assessment: Experiment 3 Interim Report (Lickteig, Sanders, Durlach, & Carnahan, in preparation).

Participants

Two groups of participants, which will be referred to as experts and novices, participated in the research. The groups differed markedly in their tactical and technical training and experience.

Experts. The expert command group participants were primarily four U.S. Army active duty LTCs. However, the participant Information Manager for the latter half of Experiment 3 runs was a major with less field and tactical experience. The major did have extensive experience with the FCS C^2 prototype as he was stationed at Fort Monmouth and assigned to the FCS C^2 operational facility. The tactical expertise of the LTCs, in particular, provided a solid base for developing and applying the technical skills required to command and control the Unit Cell assets by means of the FCS C^2 prototype interface. Their technical skills were developed in formal training sessions during each experiment, particularly Experiments 1 and 2 for comparison with novices, as described below in the Training section. Skill integration opportunities for these experts included approximately 20 experimental runs across Experiments 1 and 2, which were followed by feedback on technical and tactical performance during After Action Reviews (AARs). The extensive technical and tactical expertise amassed by the expert command group made them particularly valuable in identifying system capabilities that could be exploited by highly proficient command groups. However, the concern existed that the expert command group may not accurately portray the training needs of lesser skilled target audience command group members.

Novices. The novice command group participants were four U.S. Army cadets. The cadets selected were one junior and one senior ROTC cadet from The Ohio State University and two senior cadets from West Point Military Academy. The cadets' relatively limited tactical skills were based primarily on academic military training versus extended professional military and operational experience. In turn, the novices' relatively limited technical skills with the FCS C^2 prototype system were not grounded by tactical expertise, and did not include any formal training sessions, as described below in the Training section. Skill integration opportunities for the novices were limited to four experimental runs, and not based on relatively solid tactical and

technical skills. The novices did report having 10 to 15 years of basic computer experience, however, that included no prior experience or technical skills with any C^2 prototype system.

Experimental Design

Overall, an exploratory design that was employed stressed iterative refinement of the C^2 prototype and Unit Cell organization across a series of experimental runs. Data were collected from a Summer Study of novice command group members for comparison against existing data for an expert command group gathered during three previous experiments to yield the novice versus expert comparisons. Each run required approximately 2 to 3 hours for completion of the planning and execution phases. The scope for this report was limited to the execution phase of selected novice and expert runs.

Efforts by ARI in support of training and evaluation resulted in the respective use of *deliberate practice* and run *complexity* levels. Design for deliberate practice stressed the repetition of similar runs with feedback to ensure results were based on proficient performance. The performance feedback essential to deliberate practice included end-of-day AARs. Design goals were to help participants learn, assess, and refine the new technical skills required to operate the C^2 prototype, and the new tactical skills required to exploit the Unit Cell's progressively automated assets. The execution phase for each run was limited to approximately 60-90 minutes, and the operational setting and basic mission of the Unit Cell were relatively constant across experiments, supporting deliberate practice.

Run complexity was manipulated as a three-level independent variable to investigate how changes in operational conditions impact FCS units, particularly command group performance, and to gauge the performance limits of the proposed Unit Cell organization. It was hypothesized that changing the level of run complexity would result in a corresponding change in workload, performance success, and human-system integration requirements. The experimental design varied run complexity levels as a function of METT-TC (Mission, Enemy, Troops, Terrain, Time, Civilians) for Experiments 1-3. Three complexity levels (Medium, High and Too High) were varied by increasing enemy force activity and size, eliminating a key friendly asset, and inserting civilians on the battlefield.

The setting for battle vignette runs was a virtual environment simulation of desert terrain from the National Training Center (NTC) in which the Unit Cell conducted deliberate attack missions against a conventionally equipped threat battalion (minus/plus) to clear passage lanes for a follow-on friendly force. In general, the experimental schedule called for two experimental runs to be recorded each day with a total of thirty six runs planned for Experiments 1-3 with experts, and four runs planned for the Summer Experiment with novices.

Apparatus

The FCS C^2 designers and engineers developed the hardware and software apparatus, called the C^2 prototype system. The workstations for the command group personnel, located in the surrogate C^2 vehicle, allowed them to command and control Unit Cell ground and air robotic platforms, as well as receive information from a variety of ground and air sensors. In addition,

the C2 display depicted a real-time common operating picture of the Unit Cell's battlefield situation that the command group shared with surrogate Higher echelons.

Figure 3 depicts the C^2 prototype tactical display used in the present series of experiments with the Battlefield Assistant, Enemy Intel (Image Viewer), Target Catalog, and Resource Availability windows opened. These and numerous other information windows could be opened, closed, and modified by command group personnel during the course of a run or mission.

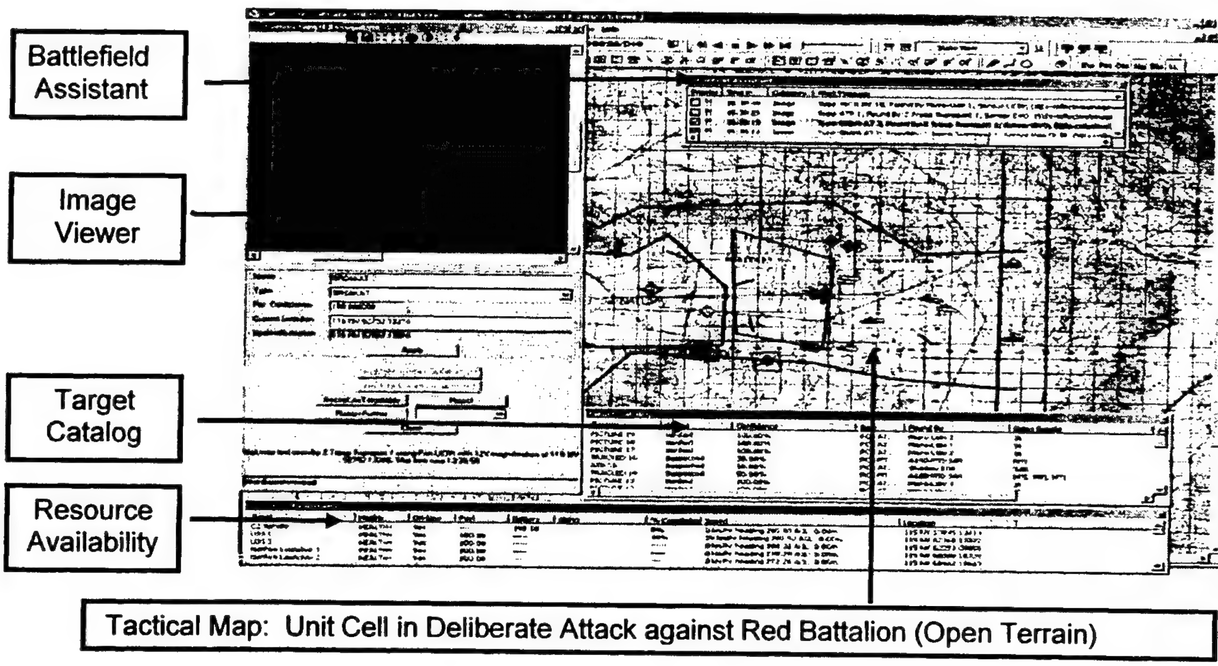

Figure 3. Command and control (C^2) prototype tactical display.

The Battlefield Assistant contains information regarding various alerts that the command group participant could specify and activate. These alerts could be set up to report whether a vehicle or the ground was hit by fired munitions. The Battlefield Assistant could also signal the command group personnel when a friendly, enemy, or neutral vehicle crossed a phase line or entered a named area of interest (NAI). The Enemy Intel window displays target images for the purpose of human target recognition (HTR), and also for battle damage assessment (BDA). The Enemy Intel window allowed the members of the command group to change a friendly or target icon's affiliation (e.g., friendly, enemy, neutral), type (e.g., tank, personnel carrier), or status (e.g., suspected target, dead target). The Target Catalog could specify available enemy targets, what friendly sensor identified the target, when the target was identified, and also the course and location of the enemy target. The Resource Availability Window could specify the health, available fuel, speed, and location of friendly Unit Cell assets.

Allocation of Team Functions. Command group personnel were given instructions to divide the command and control duties among the Commander and three subordinate command group duty positions. For the experts, the Commander decided to separate duty positions into three battlestaff managers to perform the C^2 functions of Plan, See, Move, and Strike. The Commander was responsible for the Plan function and described his duties as developing,

preparing, and synchronizing multiple courses of action (COA) to complete the mission. The Battlespace Manager was responsible for the movement of all robotic ground platforms, and the firing of the Line-of-Sight (LOS) vehicle. The Information Manager controlled the movement of all robotic aerial sensors for the purpose of reconnaissance, HTR, and BDA. The Effects Manager was responsible for the Strike function, and his duties included the control of the robotic beyond line-of-sight (BLOS) Netfire precision artillery system vehicles.

The novice command group personnel adopted an alternative strategy for allocating roles and responsibilities across duty positions, creating duty positions that were more independent and autonomous. Two of the four novice duty positions, the Commander and Information Manager, were directly modeled after the experts. However, the remaining two novice duty positions were distinguished by more decentralized function and terrain responsibilities. New duty positions were created for a North Ground Commander, and a South Ground Commander, who each controlled the Move and Strike capabilities for one Robotic sensor, FW vehicle, and LOS weapon system in the Northern and Southern sectors of the battlefield.

Training. Training primarily addressed each participant's individual operator skills with the C^2 prototype system. This training was intentionally not duty specific, but rather designed to provide cross duty skills required to operate the C^2 prototype from any of the four command group duty positions. The experts received three days of training, of which the first two were designed to address individual operator skills. On the third day of training, the experts received unstructured collective training that took place in the surrogate C^2 vehicle linked by virtual simulation to support run rehearsals. The novices' tactical training was based primarily on their limited but intense professional training as cadets in military schools and organizations, as described in the preceding Participants section. Technical skills for the novices included extensive background experience with personal computers and particularly computer-based games and simulations, as also described in the Participants section. However, novice training and skills with the C^2 prototype varied substantially. Three of the cadets had spent several weeks at the experimental site operating the C^2 system and participating in numerous runs and demonstrations with the C^2 prototype. However, one of the more senior cadets arrived late to the site and received very limited technical training on the C^2 prototype.

Measures

Measurement methods for human performance included both objective and subjective measures. Objective measures included the analysis of verbal communications and human-computer interaction. Subjective measures included surveys and questionnaires designed to measure workload, performance success, C^2 prototype effectiveness, training adequacy, and human system integration issues.

Verbal Interaction

Verbal communications were transcribed and coded using the categories of source, function, type, and factor to allow a direct comparison of the Summer Experiment and Experiment 2 verbal communications performance. A copy of the verbal communications coding scheme is provided as Appendix B. The categories of valence and command

considerations were developed after the completion of the analysis for Experiment 2. Therefore, valence and command considerations data provided by the experts in Experiment 3 were used for comparison with the findings of the Summer Experiment.

Verbal Communications were reviewed and transcribed based on quadraplex video recordings from cameras positioned on the ceiling of the C^2 vehicle across from each command group participant. This quad video image of the four primary participants in the command group greatly assisted the verbal transcription team for Experiments 2 and 3, but was unfortunately unavailable for the Summer Experiment. Researchers, therefore, utilized a digital video recording of the Commander's left display to acquire the needed audio data to prepare the transcription. Due to the extensive time and labor required to transcribe, chunk and rate all communications within a given run, verbal communications data was analyzed for only one run from each experiment.

Initially, all verbal communications were converted into written transcripts that were appended with data on source and time of communications. These transcripts were subsequently "chunked." That is, the flow of communication was blocked into units amenable for subsequent coding. This chunking of the transcript required a researcher to evaluate the transcript and then to group dialog passages together that appeared to be unitary and consistent. The goal of chunking was to create coherent blocks of dialog that were specific enough in that they did not fall under multiple rating categories.

Chunking and classifying communications into meaningful categories often requires a degree of interpretation. When different coders fail to agree, then one must question the reliability of the results. One method used to increase inter-coder agreement was to separate verbal communications transcripts into small chunks or samples of verbal data in order to narrow the range of interpretation. Inter-coder agreement across all coding categories for Experiment 2 reached 93.2%. Based on that success, small verbal communications chunks were also used when coding the verbal communications from the Summer Experiment and Experiment 3 data.

Human-Computer Interaction

The method for analyzing human-computer interactions (HCI) involved reviewing, coding, and quantifying data from digital video recordings of each of the four command group participant's two displays. Run execution video for each of the eight command group displays were reviewed by researchers for each experiment. A quality versus quantity analytic strategy focused on a relatively comprehensive analysis of HCI performance during a single run per experimental session versus a coarser analysis of multiple runs. The digital video recordings were reviewed and HCI records of the behavioral tasks performed were developed. As illustrated in Figure 4, the evaluation framework for HCI was structured to address the basic C^2 functions of Plan, See, Move, and Strike and supporting sub-function categories. The entire coding scheme is provided in Appendix D.

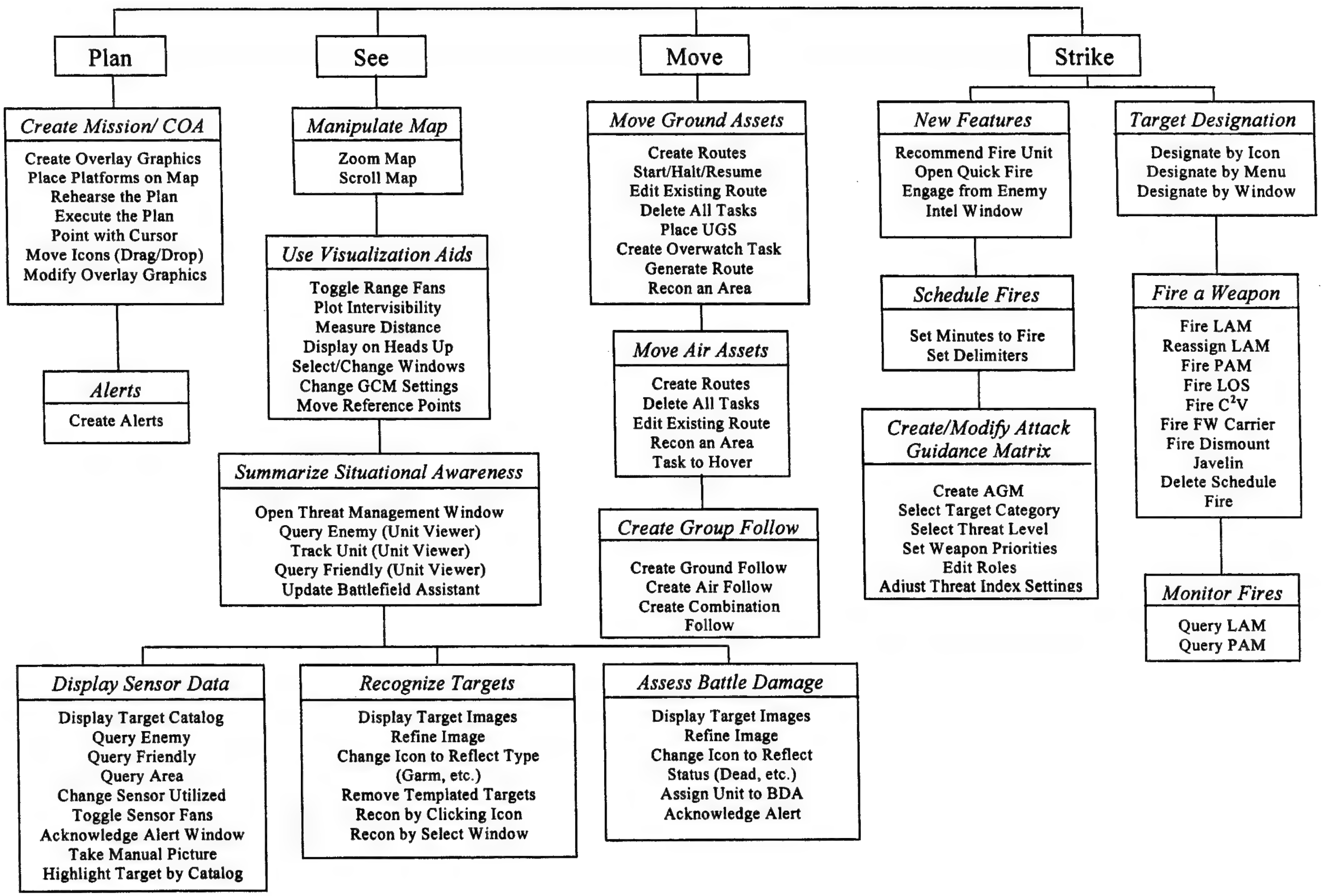

Figure 4. Human-computer interaction (HCI) coding scheme.

The HCI task performance criteria included task frequency, duration, and errors. Transcripts described each HCI task performed, recorded Start and Stop times where appropriate, and annotated any performance errors observed. For HCI tasks that routinely required 5 seconds or less to perform, such as using Zoom and Scroll Map tools, times were not recorded. However, special attention was paid to record Start and Stop times for any instances where tasks took longer than 5 seconds to perform, such as creating routes for ground or air robotic platforms.

Two researchers independently rated one of the HCI task records from Experiment 2, and ratings were compared to generate an estimate of inter-rater reliability. The independent ratings comparison yielded an index of agreement between raters of 99%. One of the original Experiment 2 researchers rated each of the eight video recordings from the Summer Experiment Run 4.

Subjective Measures

A number of subjective measures were employed by ARI during Experiment 2, Experiment 3, and the Summer Experiment. These measures included an assessment of workload and performance success, C^2 prototype support of FCS C^2 functions and METT-TC factors, training adequacy, and human systems integration. A copy of each subjective measure utilized is provided in Appendix E.

Workload and performance success. For both novices and experts, immediately after the command group participants completed an experimental run, the participants exited the C^2 vehicle and completed a brief survey on workload and performance success. Participants rated their perceived Workload across five dimensions: mental, physical, temporal, effort, and frustration (1 = Low to 100 = High). The workload questions and dimensions were adapted from the Task Load Index (TLX) developed by National Aeronautics and Space Administration (NASA)—Ames Research Center (1986). Performance success was rated on this same questionnaire (1 = Failure to 100 = Perfect).

C^2 prototype support of FCS C^2 functions and METT-TC factors. The novice participants completed this survey after the first day of experimental runs, and the experts completed the survey after the final run of Experiment 2. The participants provided estimates of the C^2 prototype's effectiveness for C^2 functions (Plan, See, Move, Strike) and METT-TC factors (Mission, Enemy, Troops, Terrain, Time, Civilian Considerations).

Training adequacy. A structured interview was conducted to investigate training issues, and to have the command group participants describe their duty position in their own words. This survey was conducted for the novices at the completion of the first day of runs. The experts completed the survey at the end of the three day training session, and also after the last run.

Human systems integration. The ARI effort to improve measurement methods included adaptation of an instrument developed by the Army Research Laboratory (ARL) to assess Apache Longbow Helicopter (AH-64D) crew stations (Durbin, 2002). This Human Systems Integration Questionnaire (HSIQ) addressed two basic areas: task workload and C^2 prototype usability.

Results

The results of the novice versus expert comparisons for objective and subjective data are presented in this section. Reported results describe both significant and non-significant differences for objective measures, and provide descriptive statistics for subjective measures. Overall, many significant and interesting differences between the novice group and the expert group are reported. Although findings are preliminary, the differences reported provide useful indicators for distinguishing novice and expert performance, with implications for training design. The results, however, must be interpreted with caution. It is important to realize that the differences reported are based on data from a single expert group and a single novice group. A significant result might simply reflect a difference between the two participant groups rather than between the populations of experts and novices they represent.

Verbal Communications

This section presents the results of comparisons between the performance of expert and novice groups during battle execution runs based on the percentage of command group verbal communications devoted to battle command functions and specific types of communication. First, a brief review of experimental run characteristics is provided. Next, results are provided for: communication silence, communication by source, function, type, METT-TC, valence, and command considerations. Supplemental figures provided for each section are based on percentage data to account for the differences in run times for the novice and expert command groups.

Run Characteristics. Table 6 summarizes key characteristics of the experimental runs analyzed to compare novices and experts. The run characteristics include duration, cumulative silence time, and number of verbal chunks. Notably, the novice's experimental run lasted approximately 1 hour and the expert's experimental run lasted approximately 1.5 hours. This difference in run time is the primary reason comparisons are based on percentage data. Source, function, type, and METT-TC factors and sub-factors comparisons are made using expert Experiment 2 data. Valence and command consideration comparisons are made using expert Experiment 3 data.

Table 1

Characteristics of the Analyzed Runs for Verbal Analysis

Expertise Level	Novice	Expert	Expert
Experiment Identifier	Summer Experiment	Experiment 2	Experiment 3
Complexity	High	High	High
Run Number	4	10	11
Run Duration	54 minutes	87 minutes	84 minutes
Cumulative Silence	7 min, 6 sec	2 min 18 sec	2 min 54 sec
Number of Verbal Chunks	245	494	461

The novice command group spent considerably more time in silence during run execution compared to the expert command group. While cumulative time in silence was under 3% during Run 10 of Experiment 2 and 4% for Run 11 of Experiment 3 for the expert command group, cumulative time in silence was over 13% for the novice command group. Possible explanations for this difference include differences in duty position roles, prior run experience, and command group cohesion. Recall, the novice duty positions fostered more autonomy and, therefore, required less discussion and collaboration before acting. The experts' command group structure required collaboration with and approval by the Commander before acting.

Communication by Source. A source code was used for verbal behavior analysis to identify who was speaking to whom. Figure 5 displays the percentage of command group verbal communications broken out by expertise group (Novice/Expert) and source (Within Cell, Cell/Subordinate, and Cell/Higher). For novices and experts, verbalizations were predominantly within the Unit Cell. Communication with the Higher echelon was significantly greater for the novices [x^2 (1, N = 739) = 9.042 p = .003]. From closer examination of the verbal transcripts communications by Higher generally involved guiding novice reconnaissance efforts and ensuring target identification before target engagement. No differences between the novices and experts were found in the percentage of communications for the Within Cell and Cell/Subordinate source categories.

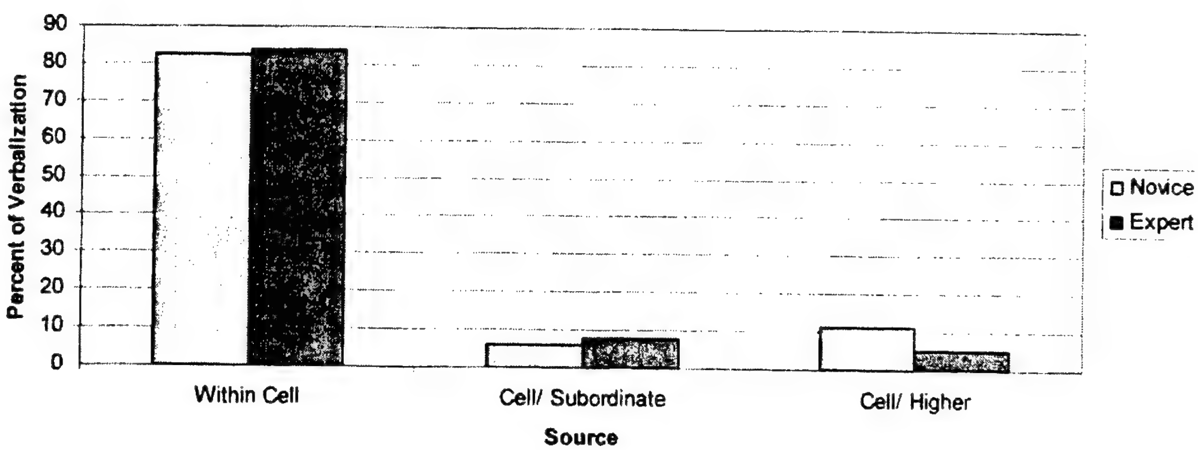

Figure 5. Percent of verbalization by Source and Expertise group.

Communication by Function. Figure 6 displays the percent of verbalizations by C^2 function and expertise level. Communications related to the Strike and Plan functions were the most frequent for the novices, 30% and 28% respectively. Communications related to the Plan and See functions were the most frequent for the expert participants, 30% and 24% respectively. Communications supporting the See function were performed less by the novices compared to the expert group [x^2 (1, N = 739) = 9.79, p = .002]. However, Strike related communications were performed more by the novices [x^2 (1, N = 739) = 9.283, p = .002]. The novices' lack of experience and training may have limited their abilities to "see" the battlefield and assess the multitude of threat images presented by the C^2 prototype. In contrast, the Strike related communications by the novices indicated their concern with destroying the enemy force. It should be noted that many of the novices' Strike engagements, however, were performed without

proper identification of enemy targets as indicated by Higher Headquarter's repeated warnings to ensure target identification before firing on targets.

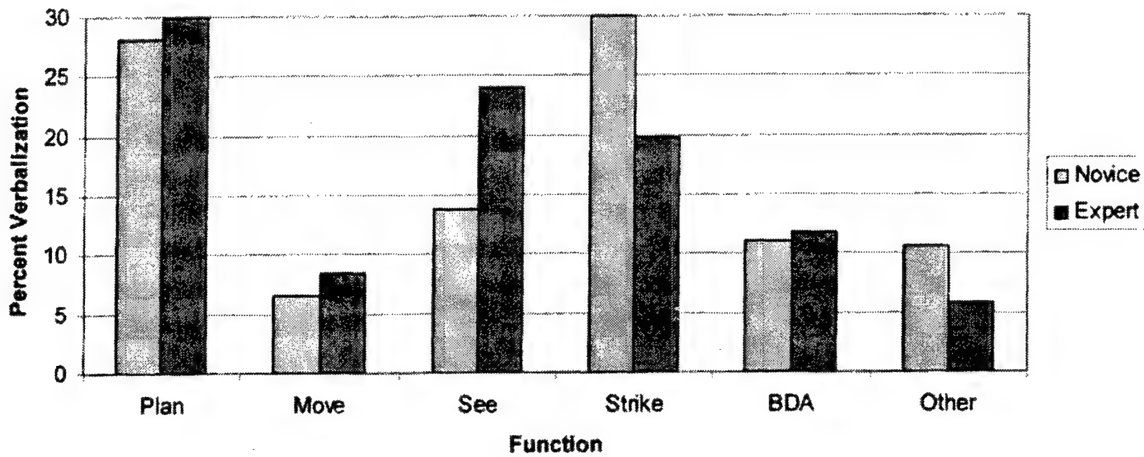

Figure 6. Percent of verbalization by Function and Expertise group.

Communication by Type. Communications were categorized by type to help identify the purpose of the command groups' communications. Figure 7 depicts the percentage of verbal communications by type for the novice and expert command groups. The distributions by type are fairly similar across both expertise levels. Share and Ask categories accounted for almost 70% of all verbalizations for both expertise levels. The "decide" category of verbalizations which involved discussion of a decision or plan were significantly lower for the novices [x^2 (1, N = 739) = 9.190, p = .002].

Why did the novices discuss decisions less? One possible explanation is that the North and South Commanders felt more independent and were less concerned about getting approval from the Unit Cell Commander. Another possible explanation may be that the novices displayed minimal discussion of target identification and were more concerned with destroying targets as soon as they were located by the C^2 system.

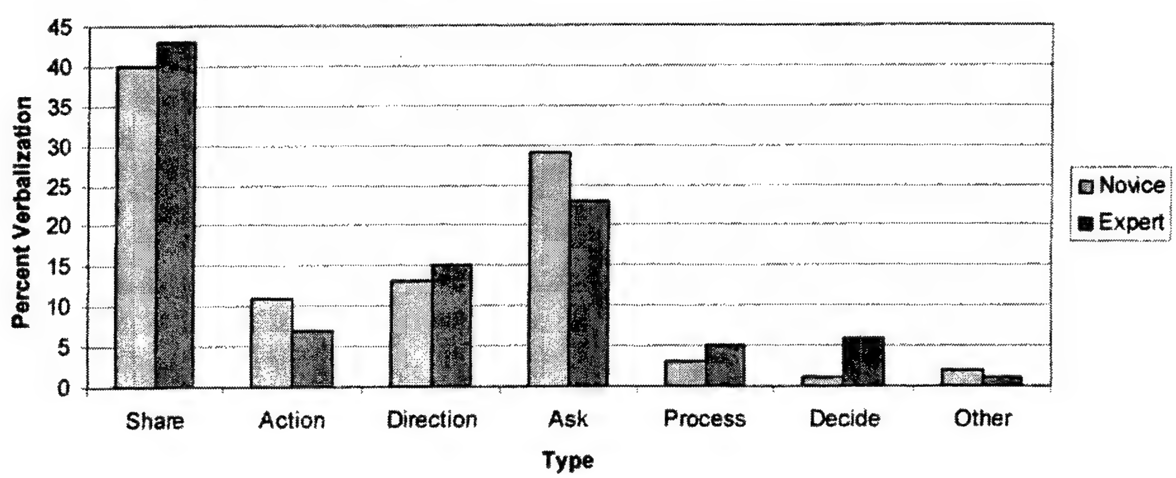

Figure 7. Percent of verbalization by Type and Expertise group.

Communication by METT-TC Factors. Figure 8 shows the frequency of verbal communications by METT-TC factors by expertise. Troops and Enemy related communications encompassed most of the verbalizations, accounting for approximately 90% of all verbalizations by novices and experts. Of the six METT-TC factors, two were found to be significantly different by expertise: Enemy and Troops. The novices made fewer Enemy related verbalizations [x^2 (1, N = 739) = 4.372, p = .037], and more Troops related verbalizations [x^2 (1, N = 739) = 4.831, p = .028] compared to the expert group. Civilian considerations were not discussed by either the novice or expert command group participants.

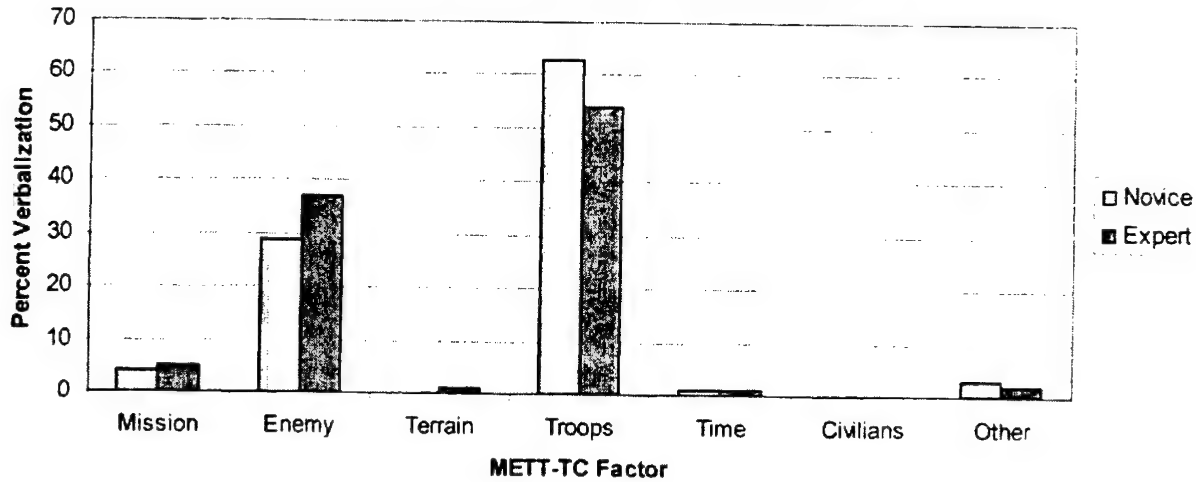

Figure 8. Percent of verbalization by METT-TC Factor and Expertise group.

The METT-TC factors were further divided into 25 sub-factors, as defined in the Verbal Coding Scheme located in Appendix B. For example, Enemy categories included four sub-factors: Location, Identification, Disposition, and Battle Damage Assessment (BDA). Figure 9 shows the distribution of communications by Enemy sub-factors and expertise level. For the novices, almost 90% of Enemy verbalizations addressed the location of enemy units and BDA. For the experts, the primary concern was the identification of the enemy units. The percentage of communications for each sub-factor, except BDA, was found to differ between the two expertise groups significantly. Location related verbalizations by novices exceeded the experts' [x^2 (1, N = 252) = 14.391, p < .001). However, the percentage of novice verbalizations was significantly less for Identification [x^2 (1, N = 252) = 12.824, p < .001] and Disposition [x^2 (1, N = 252) = 6.178, p = .013] sub-categories compared to the expert group.

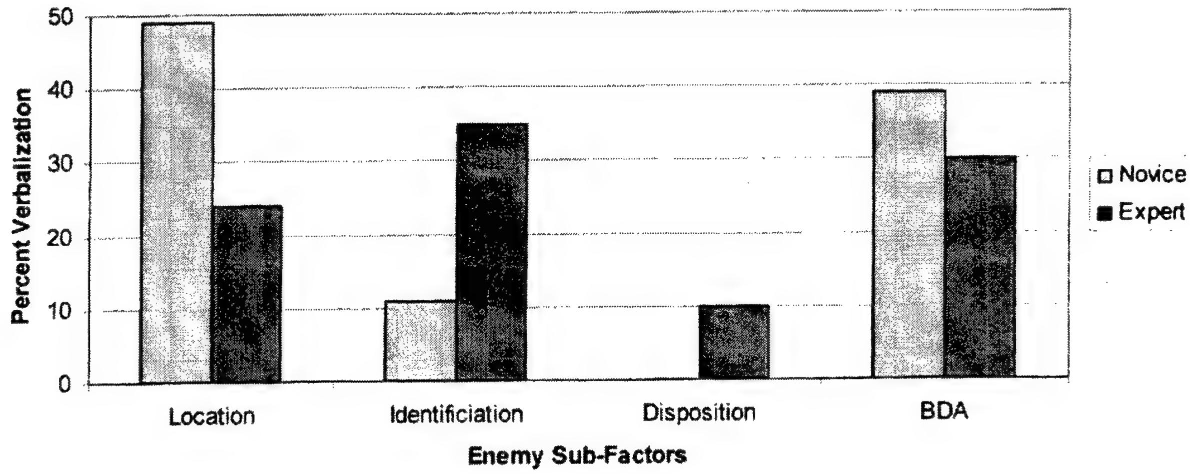

Figure 9. Percent of verbalization by Enemy Sub-factors and Expertise group.

Similarly, Troops related communications were partitioned into fourteen sub-factors, as shown in Figure 10. For both the novices and the experts, the Strike-Lethal sub-factor dominated the communications, which is related to launching, firing, and deploying weapons with intent to destroy (e.g., the employment of Precision Attack Missiles, referred to as PAMs). Of the fourteen sub-factors, three were found to differ significantly across the expertise groups. The novice group devoted significantly greater percentage of their communications to discussing Loss sub-factor issues compared to the expert group [x^2 (1, N = 421) = 4.862, p = .027]. In contrast, the expert group devoted a greater percentage of their communication to the discussion of Mobility issues [x^2 (1, N = 421) = 21.014, p < .001], and sensors [x^2 (1, N = 421) = 5.308, p = .021].

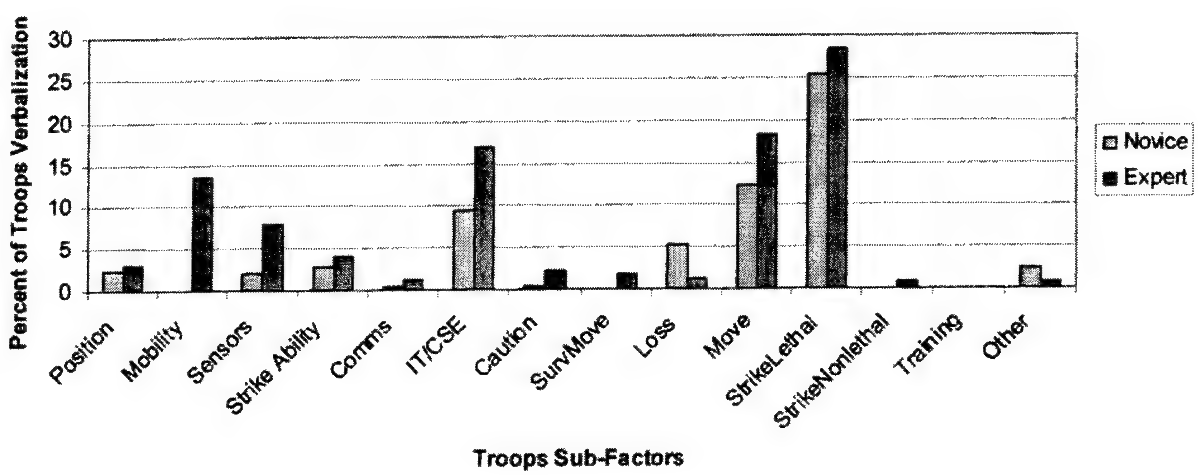

Figure 10. Percent of verbalization by Troops Sub-factors and Expertise group.

Communication by Valence. The communications transcript chunks were also coded for valence to indicate whether verbal communications represented negative, neutral, or positive information. See Appendix C for an example of Valence codes assigned to verbal chunks from

Experiment 3. Results on valence by function and expertise are summarized in Figure 11. The majority of communications for both novices and experts were positive for Plan, See, Move, and Strike communications. However, BDA verbalizations were communicated more negatively by the novices and more positively by the experts. As a result, only the comparisons for BDA were significantly different. Novices communicated less positively [x^2 (1, N = 87) = 5.882, p = .015] and more negatively [x^2 (1. N = 87) = 8.884, p = .003] about BDA. The novices were constantly performing BDA image analysis during the Summer Experiment, and were frequently frustrated by the lack of damage inflicted upon target images. Their frustrations may have been eased if they had devoted more time to target identification, which may have influenced them to use munitions capable of destroying heavily armored targets.

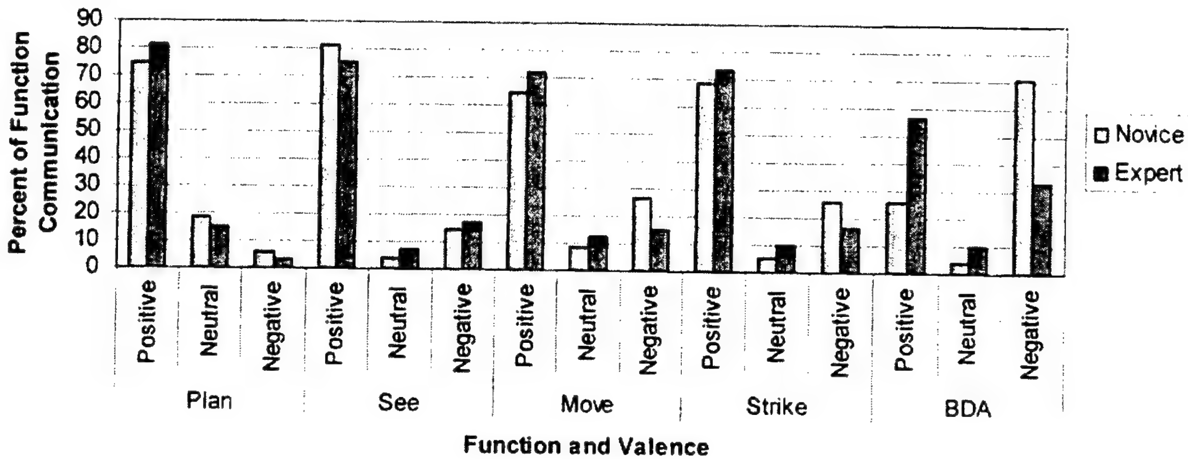

Figure 11. Percent of verbalization by Function, Valence, and Expertise group.

Command Considerations. The command group discussions were analyzed to identify instances of key battle command considerations. The command considerations coding provides a more explicit linkage indicating how the verbal behaviors of participants relate to the cognitive processes required for battle command. Table 2 contains the nine topics used to assess command considerations.

Table 2

Command Considerations for Analyzing Command Group Verbalization

Plan: Execution is self-initiated and preceded by plan coordination/refinement.
Inform: Make information requirements known.
See: Battlefield visualizations that are dynamic/predictive/proactive.
Coordinate: Create synergistic effects with multiple assets/teamwork.
Assets: Use all assets available.
Situation Awareness: Continual situation assessment, dynamic/contingency planning.
Terrain/Time: Consider effects of terrain/time.
Enemy: Model a thinking enemy.
Mission: Keep sight of the big picture and mission intent.

Figure 12 illustrates the percent of communications for each of the nine categories of command considerations. All nine command considerations were communicated by the experts, but only six were communicated by the novices. The overall frequency of command considerations totaled 13 of 245 total communication chunks for novices and 54 of 461 total communication chunks for experts. Two of the nine command considerations were found to be significantly different by expertise. Significant differences occurred where there was no communication in support of that command consideration by the novices with a relatively higher level verbalized by the experts. The Plan command consideration refers to the execution of events being self-initiated and preceded by a plan coordination or refinement. Novices communicated a significantly smaller percentage of Plan related command considerations (p = .031, two-tailed). The See command consideration includes communications concerning battlefield visualizations that are dynamic, predictive, and/or proactive. Novices also communicated a significantly smaller percentage of See related command considerations (p = .031, two-tailed) compared to the expert command group.

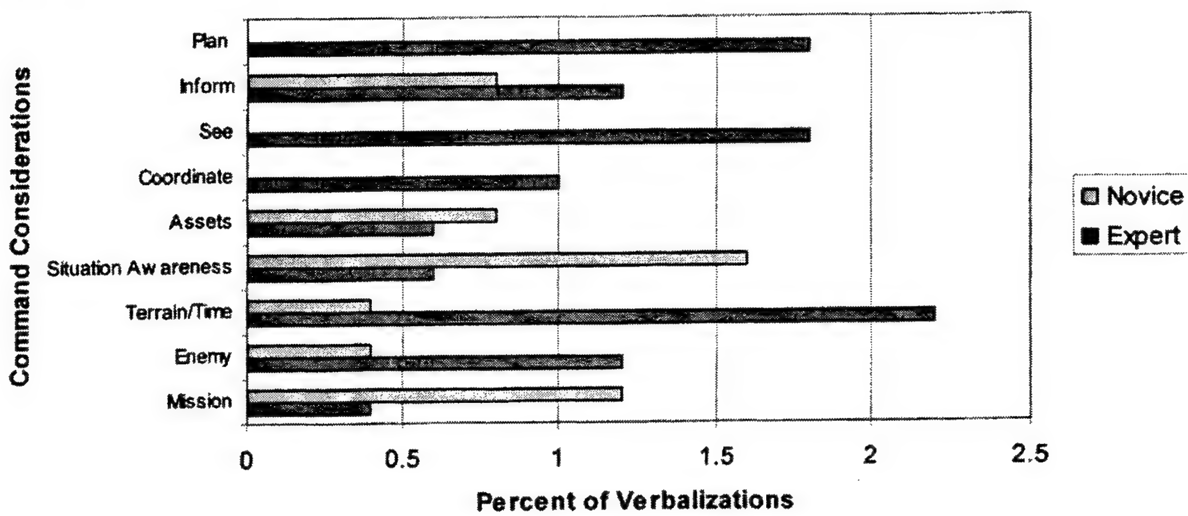

Figure 12. Percent of verbalization by Command Considerations and Expertise group.
Note: Command Considerations are defined in Table 2.

Human-Computer Interactions (HCI)

The HCI results provide a detailed account of the interactions required to command and control the Unit Cell. Overall, results indicated that the pattern of human-computer interactions differed substantially between novices and experts. The primary measures of HCI performance were frequency of performance, and performance time. Performance errors were less useful as performance criteria because very few errors occurred. While performance times were collected for selected long duration tasks, performance times did not appear to be associated with any indicator of task performance success or failure. Findings are limited by comparisons based on only one run per expertise level.

Time Duration for HCI. Time duration for completion of an interaction was examined as an important aspect of HCI performance. Interactions requiring longer performance times could be candidates for redesign, machine aiding, or additional training, especially if these times differ

greatly between novice and expert users. Performance time was estimated by identifying Start and Stop actions, typically involving the selection of a menu option. Frequency of occurrence and mean performance times for the novice and expert users are provided in Table 3 for interactions that required more than five seconds to perform, referred to as "long duration" tasks. Approximately 15.5% of all interactions performed by the novices (162 of 1,047) required more than five seconds to perform, and 17.0% of the interactions performed by the experts (234 of 1,376) required more than five seconds, a difference between the novices and experts that is non-significant.

Significance tests were performed on the individual long duration interactions to determine if the novices and experts performed the interactions at different mean levels or if one group's task performance was more varied. The novices were significantly slower in creating ground routes [$t(55) = 2.316$, $p = .024$], and were significantly more varied in their performance of creating air routes than the expert participants ($F = 15.663$, $p < .001$). The novices' performance limitations in creating ground routes and increased variance in creating air routes may indicate that the novice command group technical skills were limited by their lack of training and experience with the C^2 prototype. The experts were significantly more varied in their performance of recognizing targets (HTR) than the novices ($F = 15.277$, $p < .001$). This significant difference may have been caused by the higher quantity of interactions related to HTR performed by the experts than the novices.

The experts were also significantly more varied in their performance of rebooting their display after a system crash ($F = 9.330$, $p = .004$). This significant difference may be explained by anecdotal evidence that the novices and experts partitioned the functionality of the C^2 prototype differently on their individual dual displays. The novices used their dual displays to present different information and C^2 system functionality. Therefore, when one of novices' displays crashed, they were forced to quickly restart the system in order to access the information and functionality lost by the crash. The experts, however, set up their displays to present similar, if not redundant, information and C^2 system functionality. This similarity of display setup allowed the experts to rapidly switch attention to the functioning display, and in many cases, system designers remotely restarted the expert's crashed display. No significant differences were found between the expert and novice command groups for the Assess Battle Damage task.

Table 3

HCI Long* Duration Interaction Frequency and Time

Long Duration Interaction	Frequency of Interaction		Time in Seconds (Mean/SD)	
	Novices	Experts	Novices	Experts
Create Ground Route	26	31	15.4/12.8	9.7/7.8
Create Air Route	23	50	17.4/17.5	12.1/4.7
Recognize Targets	23	77	13.1/7.9	17.2/19.6
Assess Battle Damage	75	37	12.0/7.5	12.3/9.8
Crash Reboot	15	39	34.3/8.8	43.5/41.6

*Long duration interactions require more than five (5) seconds.

The HCI by Function. Figure 13 presents the percentage of interactions performed by novices and experts across See, Move, and Strike functions. The novices in the Summer Experiment and experts in Experiment 2 utilized the same version of the C^2 prototype's software which contained the exact same functionality and system tools. No Plan interactions were recorded during the execution phase of the Summer Experiment Run 4, or Experiment 2, Run 10. The Plan related interactions would be more likely to occur during the planning phase of each run.

For both novices and experts, the greatest percentage of interactions performed was devoted to the See function, versus Strike and Move. No differences in percentage were found for the Move function; however, a significant difference was discovered for See and Strike. Novices performed fewer See related interactions [x^2 (1, N = 2423) = 4.725, p = .030], and devoted a higher percentage of their total interactions to the Strike function [x^2 (1, N = 2423) = 20.455, p < .001], which corresponds to the pattern of novice versus expert verbalizations for See and Strike.

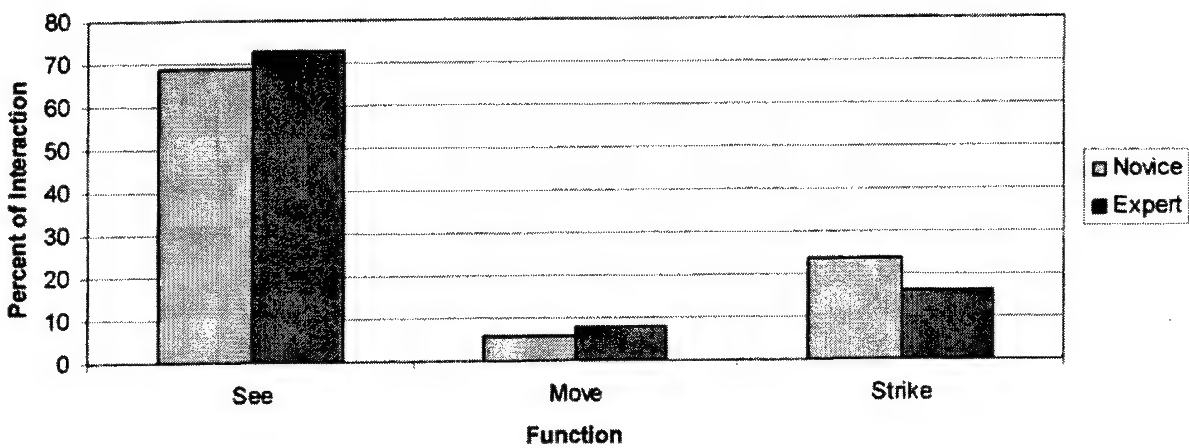

Figure 13. Percent of human computer interactions (HCI) by Function and Expertise group.

The HCI by Function, Sub-Function, and Duty Position. Table 4a and 4b provide frequency data by C^2 function and the ten sub-functions performed by the novices and experts during the runs analyzed. Of the 17 HCI sub-functions previously depicted in Figure 4, 10 were performed by both the novices and experts. The tables reflect the differences in duty position activities assumed by the novices and experts.

Table 4a

Novice Frequency and Percent of human-computer interaction (HCI) by Duty Position

Function Sub-Function	Commander		South Commander		Information Manager		North Commander	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
See	206	90.0	112	42.1	304	91.8	98	44.3
Manipulate Map	66	28.8	22	8.3	41	12.4	11	5.0
Use Visualization Aids	17	7.4	12	4.5	12	3.6	24	10.9
Display Sensor Data	29	12.7	74	27.8	22	6.6	30	13.6
Recognize Targets	43	18.8	2	0.8	94	28.4	7	3.2
Assess Battle Damage	51	22.3	2	0.8	135	40.8	26	11.8
Move	9	3.9	14	5.3	24	7.3	15	6.8
Move Ground Assets	4	1.7	14	5.3	3	0.9	15	6.8
Move Air Assets	5	2.2	---	---	21	6.3	---	---
Strike	9	3.9	139	52.3	2	0.6	100	45.2
Designate Target	2	0.9	61	22.9	1	0.3	46	20.8
Fire a Weapon	3	1.3	64	24.1	1	0.3	51	23.1
Monitor Fires	4	1.7	14	5.3	---	---	3	1.4
Other	5	2.2	1	0.3	1	0.3	8	3.6
Total	229	100	266	100	331	100	221	100

Table 4b

Expert Frequency and Percent of human-computer interaction (HCI) by Duty Position

Function Sub-Function	Commander		Battlespace Manager		Information Manager		Effects Manager	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
See	207	94.5	246	59.7	330	82.5	220	63.8
Manipulate Map	3	1.4	45	10.9	43	10.8	74	21..4
Use Visualization Aids	34	15.5	29	7.0	17	4.3	8	2.3
Display Sensor Data	128	58.4	142	34.5	42	10.5	135	39.1
Recognize Targets	24	11.0	24	5.8	178	44.5	2	0.6
Assess Battle Damage	18	8.2	6	1.5	50	12.5	---	---
Move	7	3.2	43	10.4	53	13.3	5	1.4
Move Ground Assets	7	3.2	43	10.4	3	0.8	5	1.4
Move Air Assets	---	---	---	---	50	12.5	---	---
Strike	1	0.5	107	26.0	---	---	118	34.2
Designate Target	---	---	57	13.8	---	---	37	10.7
Fire a Weapon	---	---	49	11.9	---	---	54	15.7
Monitor Fires	1	0.5	1	0.2	---	---	27	7.8
Other	4	1.8	16	3.9	17	4.3	2	0.6
Total	219	100	412	100	400	100	345	100

Figure 14 identifies the 10 sub-functions performed by the novices and experts that allow for a comparison of performance across these expertise groups. Overall, significant differences

were found between the expertise groups for five of the 10 sub-functions. Under the See function, Display Sensor Data and Assess Battle Damage sub-functions were found to be significantly different between the novices and experts. Display Sensor Data includes displaying enemy and friendly property information by cursoring over the display icon, and was performed significantly less by the novices [x^2 (1, N = 2423) = 98.605, p < .001]. This result suggests that the novices were performing relatively fewer interactions to maintain their situational awareness of the battlefield. Assess Battle Damage includes opening and viewing a target image for the purpose of detailing the extent of damage inflicted by previously fired munitions, and was performed to a greater extent by the novices [x^2 (1, N = 2423) = 127.350, p < .001]. This result suggests that the novices were more involved in determining the effect of their munitions on enemy targets.

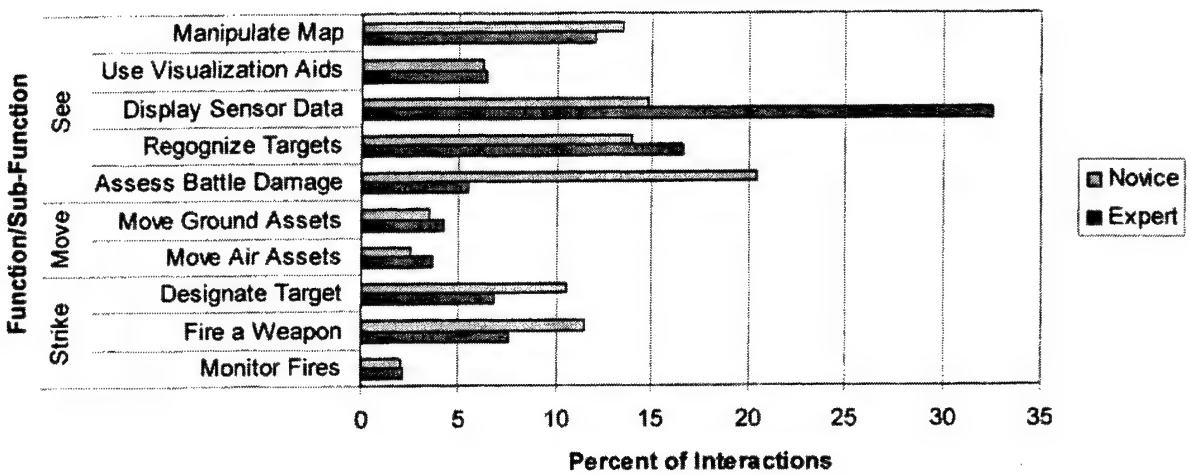

Figure 14. Percent of human-computer interaction (HCI) by Function, Sub-function and Expertise group.

Within the Move function category, the percentage of HCI actions performed to accomplish the Move Air Assets sub-function was found to be significantly higher for the experts than the novices [x^2 (1, N = 2423) = 4.326, p = .038]. In this experiment, all air assets were sensors used to support either HTR or BDA activities. The result suggests that the experts, who used the sensors to support both HTR and BDA, performed somewhat more movement of air assets than the novices, who were primarily concerned with BDA. Under the Strike function, the percentage of HCI actions devoted to Designate Target and Fire a Weapon tasks were significantly higher for the novices than the experts. Designate Target, which is performed by clicking a target icon, selecting the target from a window, or a selection from a drop down menu, was performed significantly more by the novice participants [x^2 (1, N = 2423) = 9.942, p = .002]. Fire a Weapon, the sub-function, which includes firing the Line-of-Sight (LOS) vehicle or launching a loitering attack missile/precision attack missile (LAM/PAM) from the Netfire platform, was also performed significantly more by the novices [x^2 (1, N = 2423) = 10.296, p = .001]. These results suggest that the novices' may have had a more "shoot first and ask questions later" attitude towards target identification.

Further analysis of Display Sensor Data and Assess Battle Damage interactions revealed significant differences between the novices and experts. The six categories of interactions within Display Sensor Data sub-function performed by both the novices and experts are displayed in Figure 15. The frequency of HCI actions performed within four of the six categories was found to differ significantly across the expertise groups, all of which were performed less frequently by the novices. Significance levels for these four category comparisons are presented below.

- Target Query: x^2 (1, N = 2423) = 35.391, $p < .001$
- Friendly Query: x^2 (1, N = 2423) = 18.561, $p < .001$
- Area Query: x^2 (1, N = 2423) = 6.343, $p = .012$
- Alert Window Confirmation: $p < .001$, two-tailed

A Target Query is performed by cursoring over a single target icon on the display to bring up a properties window that presents information such as enemy course and speed. A Friendly Query is performed by cursoring over an individual friendly icon on the display to bring up an identical properties window for the friendly platform. An Area Query is performed by cursoring over a collection of enemy, friendly, or a combination of platforms on the display, and to bring up a properties window that includes the information for all platforms within that area. An Alert Window Confirmation requires an acknowledgement of system alerts or warnings (e.g., Draega 23 has entered NAI Orange). The Alert Window Confirmation was performed less frequently by the novices than the experts. The novices utilized the system less in aiding their situational awareness and did not create an active alert system to signal various available system cues and warnings.

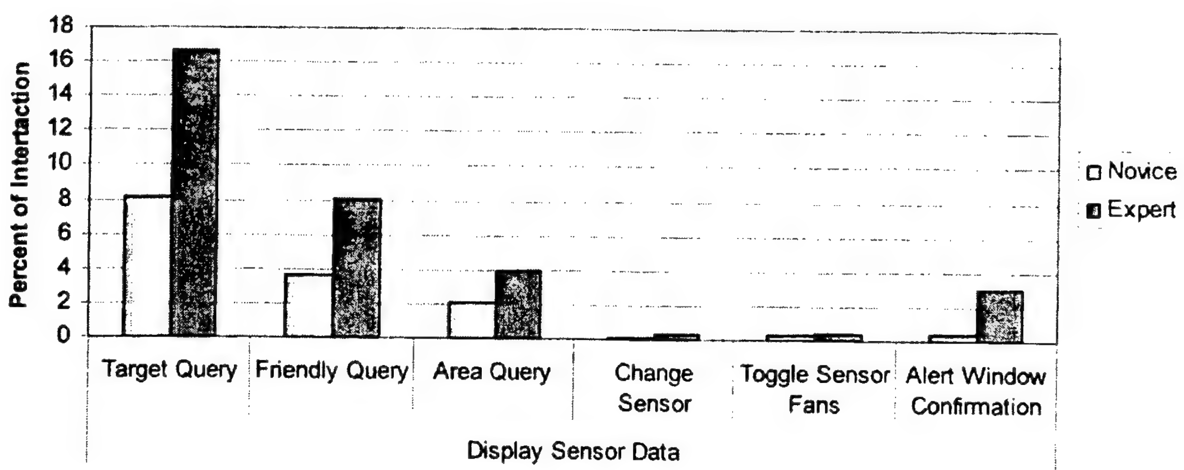

Figure 15. Percent of human-computer interaction (HCI) by Display Sensor Data Sub-function, and Expertise group.

Further analysis of the Assess Battle Damage sub-functions revealed significantly higher levels of interactions by the novices. The three sub-functions within Assess Battle Damage performed by both the novices and experts are displayed in Figure 16. The Display Target Images sub-function interactions, which require the command group participant to open the Enemy Intel window for the purpose of displaying a BDA image, were performed more frequently by the novice group compared to the experts [x^2 (1, N = 2423) = 82.854, $p < .001$].

The Refine Target Images sub-function, which includes altering the contrast and brightness of the image, and also zooming and scrolling the image for better viewing was also performed more frequently by the novices [x^2 (1, N = 2423) = 37.218, p < .001]. The Change Icon State sub-function, which includes changing the enemy's icon status from suspected to dead, was also performed more often by the novices [x^2 (1, N = 2423) = 4.743, p = .029].

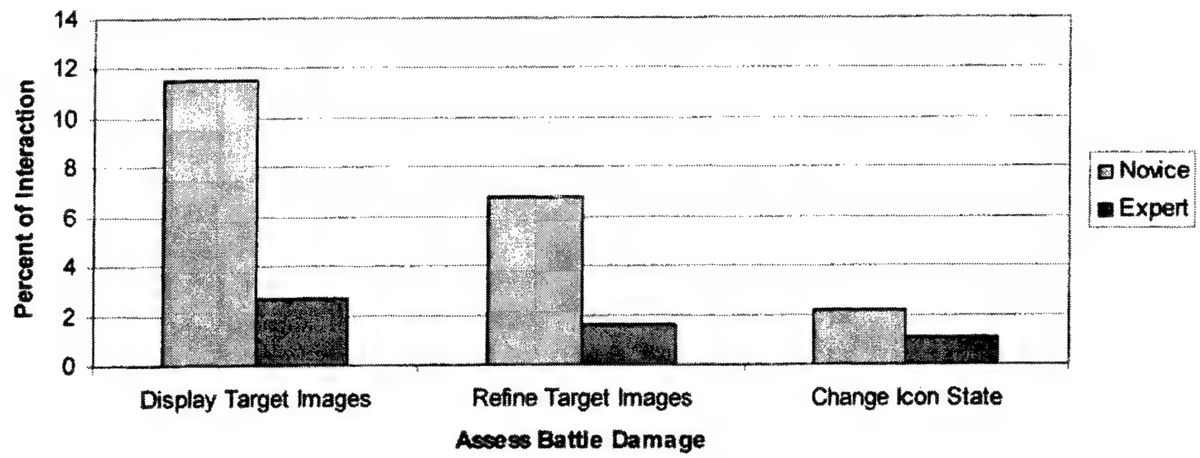

Figure 16. Percent of human-computer interaction (HCI) by Assess Battle Damage sub-function, and Expertise group.

The HTR and BDA Image Analysis. Recognize Targets and Assess Battle Damage, more commonly referred to as human target recognition (HTR) and battle damage assessment (BDA) respectively, were singled out for assessment due to their demanding human performance requirements. Figure 17 displays the percentage of images reviewed by novice and expert command group participants by type of image (HTR/BDA). This analysis revealed that the novices performed significantly less image analysis in support of HTR, and significantly more image analysis in support of BDA [x^2 (1, N = 262) = 56.953, p < .001]. These results reinforce the notion that the novices were less concerned with identifying the type of enemy targets, and more concerned with assessing the effect of fired munitions.

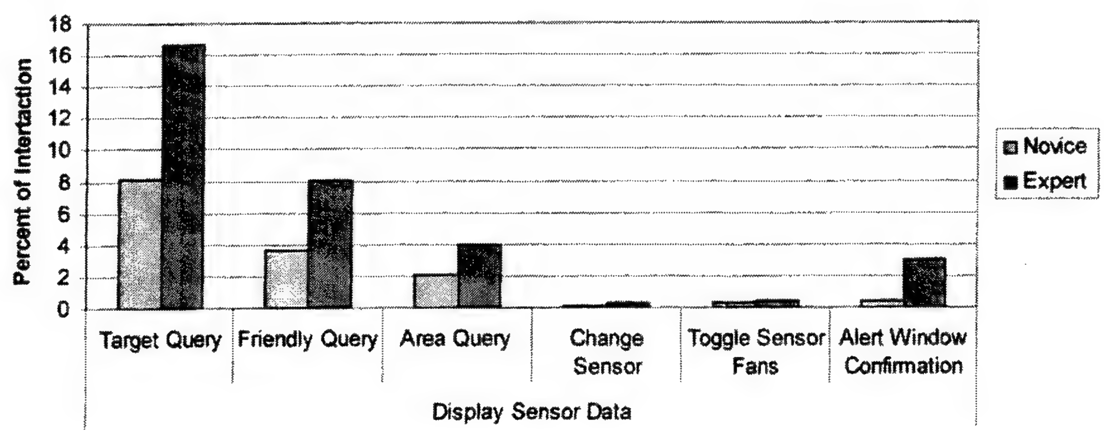

Figure 17. Percent of image analysis human-computer interaction (HCI) by Type and Expertise group.

Redundant HTR and BDA. The HTR and BDA task demands may be exacerbated by redundant work across participants. Redundant HTR and BDA performance refers to the situation where two or more command group participants examine the *same* target image at different times within a run. Multiple reviews of the same target image by different members of the command group at the same time without use of the heads up display may also indicate redundant performance. How often was the same sensor image reviewed at different times by different participants? For the novices, 46 of 150 (30.7%) total images were reviewed by more than one participant. For the experts, 69 of 112 (61.6%) of total images were reviewed by more than one participant. Notably, multiple reviews of the same images by the novices were significantly less than the experts [x^2 (1, N = 262) = 23.685, p < .001]. These differences may be explained by the fact that the novice Information Manager performed roughly two-thirds of all image analysis during the Summer Experiment Run 4, and the expert Information Manager performed roughly half of the image analysis during Experiment 2 Run 10. This difference is not statistically significant; however, the higher amount of image analysis performed by the expert Commander and remaining battlestaff managers may have contributed to the increase in redundant performance.

Subjective Measures

Results from four survey and questionnaire instruments are reported in this section: Workload and Performance Success, Effectiveness of the C^2 Prototype, Training Adequacy, and the Human Systems Integration Questionnaire. No significance tests were performed for these subjective measures.

Workload. Figure 18 displays the average workload rating across duty positions broken out by run complexity and command group expertise. Overall, the average ratings were above the scale's midpoint (50) with an average for the novices of 67.0 and an average for the experts of 61.2. For both novices and experts, there is a consistent trend of increased workload as run complexity increased. Results suggest that the manipulation of run complexity increased workload as predicted.

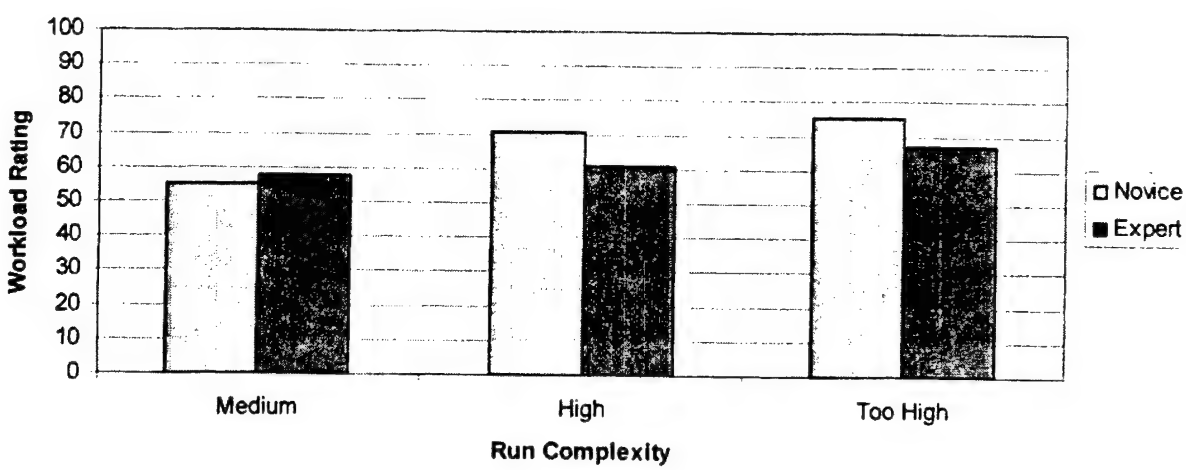

Figure 18. Average workload ratings by Run Complexity and Expertise group.
Note: 0 = Low Workload, 100 = High Workload

Performance success. Figure 19 presents the performance success ratings averaged across duty positions, broken out by run complexity and command group expertise. For both novices and experts no clear trend is evident to suggest that perceived performance success varies in a linear fashion with increases in run complexity. Performance ratings from novice command group members were complicated by the fact that their only Medium complexity level run was also their first run, which may have led them to underestimate their success on the first trial.

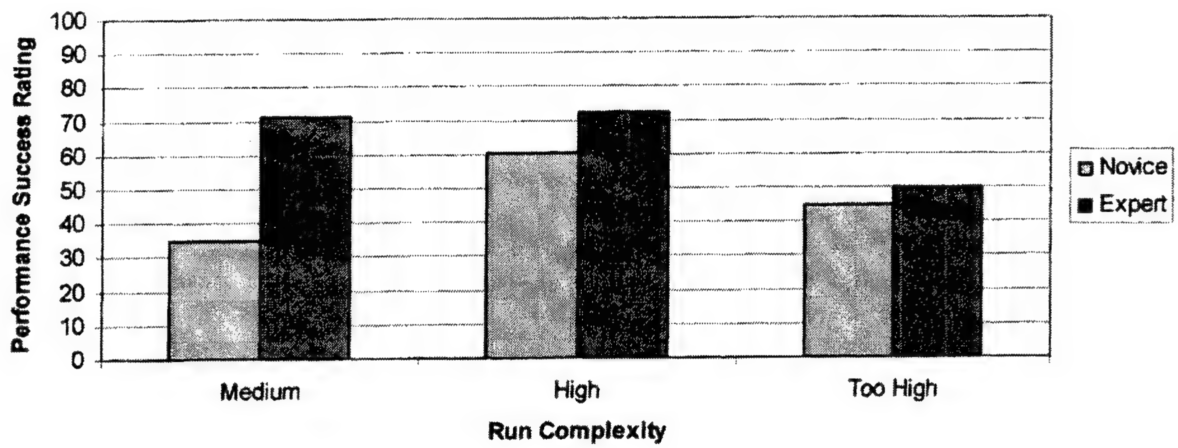

Figure 19. Average performance success ratings by Run Complexity and Expertise group.
Note: 0 = Failure, 100 = Perfect

Command and control functions. Figure 20 provides novice and expert effectiveness ratings for the C^2 prototype in support of Plan, See, Move, and Strike functions. Overall, Plan and Move functions received ratings of approximately 4 "Effective," and See and Strike functions received ratings of approximately 3 "Borderline." The effectiveness ratings were very similar across command group expertise levels, suggesting that improvements in the C^2 prototype would support both novice and expert performance.

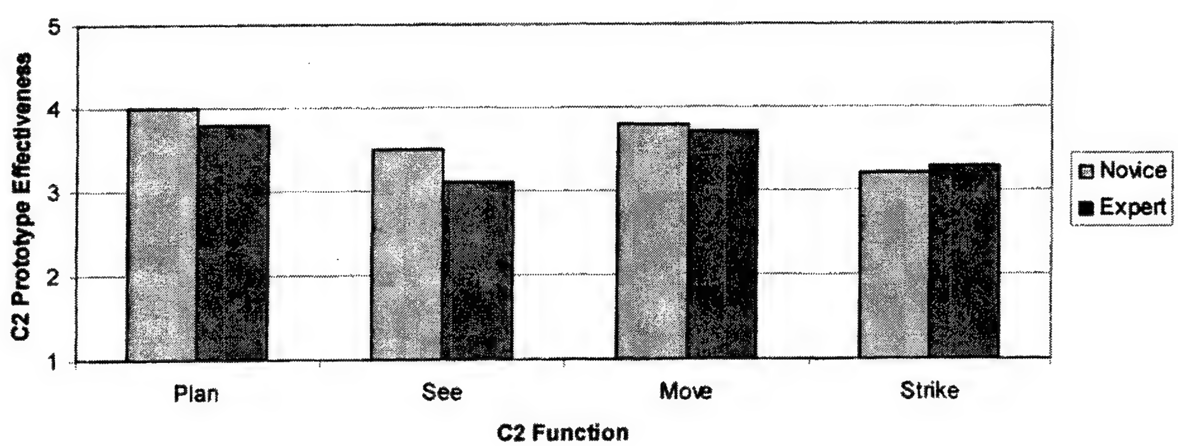

Figure 20. Mean ratings of effectiveness of command and control (C^2) prototype by Function and Expertise group.
Note: 1 = Very Ineffective, 2 = Ineffective, 3 = Borderline, 4 = Effective, 5 = Very Effective

Selected participant comments on effectiveness by function are provided below, which helped to identify the shortcomings and provide system design changes in See and Strike functions:

Selected Novice Comments:

- Need to change default settings for the search radius of the Netfire munitions.
- The number of images provided by the system can be overwhelming.
- Need to improve the images.
- Need to reduce clutter on the map display.

Selected Expert Comments:

- Need more flexible sensor search patterns.
- Need better feedback on the outcome of target engagements.
- Need more capable Aided Target Recognition (ATR).

METT-TC factors. Figure 21 displays novice and expert ratings of the C^2 prototype effectiveness in supporting METT-TC factors. Across both novices and experts, average ratings ranged between "Borderline" and "Effective." Participant comments regarding C^2 prototype effectiveness by METT-TC factor included requests for more informative and continuous intelligence feeds from higher headquarters on Enemy units, and the need to reduce the time required to task and re-task Troop assets.

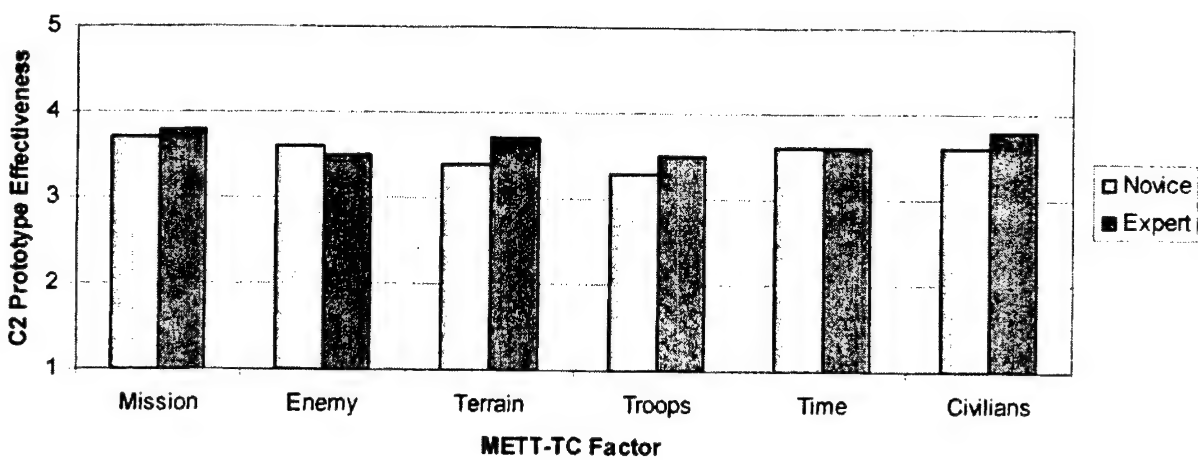

Figure 21. Mean ratings of effectiveness of command and control (C^2) prototype by METT-TC factor and Expertise group.
Note: 1 = Very Ineffective, 2 = Ineffective, 3 = Borderline, 4 = Effective, 5 = Very Effective

Training Adequacy. A structured interview was conducted to investigate training issues, and have the command group participants describe their duty position requirements in their own words. Overall, participant responses about training underscored the need for more structured, practical exercises at individual and, particularly, collective levels. Selected key comments on individual and collective training adequacy are provided below. The detailed comments suggest

that more adequate technical, tactical, individual, and collective training is needed for both novices and experts.

- Novices:
  Individual. The Cell Commander reported that individual training was not adequate. His relatively extensive experience with demonstrations and pilot testing of the C^2 prototype was primarily unstructured self-training. Lack of extended training undermined the novices' knowledge of the system's capabilities.
  Collective. Collective training was not adequate, in part due to late arrivals by some of the Cadets. A few days of practice runs may have improved command group cohesion and performance.

- Experts:
  Individual. There were many capabilities that they didn't have time to investigate or practice with. More practice would have increased trust in the system and improved the quality of the run outcomes.
  Collective. Needed more time with subject matter experts (SMEs) to increase teamwork. An increase in practice runs would improve command group cohesion.

As noted, the duty chart for the novices is located in Appendix F, and the duty chart for expert users is provided in Appendix G. Since the novices did not assume the same roles as the experts, their duty position responsibilities were not the same as the experts. The experts appeared to specialize their duty positions in terms of tasks that needed to be performed. Roughly, the duty positions were separated by function: Plan (Commander), See (Information Manager), Move (Battlespace Manager), and Strike (Effects and Battlespace Managers). The novices appeared to specialize their duty positions in terms of battlefield areas and more traditional company/platoon like organizations. The novices created a hierarchy of command, with the Unit Cell Commander and two subordinate Commanders to control equal Unit Cell assets along a Northern and Southern front. The third battlestaff manager assumed the responsibilities of the Information Manager, controlling aerial sensors, and performing image analysis across both Northern and Southern areas and Commanders.

Human Systems Integration Questionnaire. In performing the HCI analysis for Experiment 2, researchers noticed that certain tasks seemed to create more workload and frustration than others. The list of tasks examined is provided in Figure 22. The Human Systems Integration Questionnaire contained a section addressing task workload, in which the participants rated interactions on a scale from 1 ("Workload Insignificant") to 10 ("Task Abandoned").

Figure 22 displays average task workload ratings for both expertise groups. Every task was rated higher by the novices, and for ten of the seventeen tasks examined, the novices' workload ratings were double the experts' ratings. As reported in the HCI analysis, the novices task workload ratings may explain why they performed significantly fewer interactions overall, with the exception of BDA Reporting.

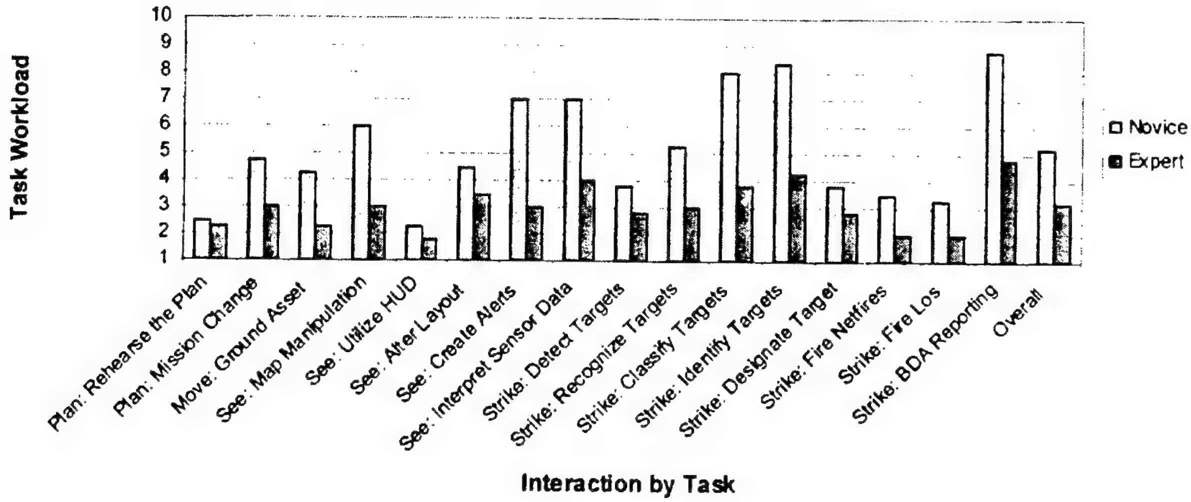

Figure 22. Average task workload ratings by Expertise group.
Note: 1 =Insignificant Workload, 10 = Task Abandoned

Ratings of the usability of the C^2 prototype were generally positive for both novices and experts. However, novices rated two display characteristics as troublesome: contrast between symbols and the background, and legibility of text on displays. Most tasks were rated by both novices and experts as being easy to perform, with only four tasks being rated as "Borderline" or above in difficulty. These four borderline tasks involve:

- Visualizing missile trajectory and intended target.
- Distinguishing between moving and stationary icons.
- Visualizing past and future threat positions.
- Determining which entity detected the threat target.

Appendix H provides the novice command group recommended changes to the default settings for the C^2 prototype, and Appendix I contains the expert group recommendations. Participant feedback in the form of "No Change" and "Recommended Change" are included for the seven default settings listed in the questionnaire and for the "Other" category. Recommendations included suggested changes to the default settings currently in place for the search radius of the LAM, ammunition type and range, and the suggestion that a provision be added to allow saving work such as image adjustments and workstation preferences. This feedback was considered by designers when improving the default settings of the C^2 prototype. In sum, the respondent feedback on the Human Systems Integration Questionnaire was used to improve operator usability of the C^2 prototype.

Discussion

Overall, the novice and expert command groups differed considerably, with the expert group holding higher levels of rank, greater background and operational experience, a longer period of experience with the FCS C^2 system, and extended personal and professional relationships. Differences between the two expertise groups were identified for performance

measures that included objective measures of verbal communications and HCI and subjective measures. All findings are regarded as preliminary given research limitations; particularly the exploratory nature of the FCS C^2 program and small participant sample size.

Tactical Skills

In the Participants section, the authors claimed that the Cadets' relatively limited tactical skills were based primarily on academic military training versus the extended professional military and operational experience of the experts. The claim of the novices' limited tactical skills was supported by the results, including verbal communications by source, function, and METT-TC factor, and HCI by function, sub-function, and individual interaction.

The verbal communications by source comparisons indicated that the novice command group devoted a significantly greater percentage of their communications to interactions with Higher than did the experts. As reported, the authors reviewed the novices' verbal transcripts to determine the purpose of the communication between the novices and Higher. Most of the communications initiated by Higher were targeted at providing tactical suggestions or reminders to the novice command group participants. These communications involved guiding novice reconnaissance efforts and ensuring target identification before target engagement.

The verbal communications by function comparisons and HCI results indicated that the novices communicated and interacted with the system less in support of: seeing the battlefield, maintaining situational awareness, and identifying targets. The novices communicated significantly less and used their C^2 prototypes significantly less to perform See related behaviors. For general situational awareness, the novices performed significantly fewer interactions in support of the Display Sensor Data sub-function.

Most importantly, the claim of the novices' tactical deficiency was supported by verbal communications and HCI results that indicated a frequent failure to identify targets before firing weapons. The novices displayed significantly fewer verbalizations in support of the identification of enemy targets sub-factor within METT-TC. The novices also performed significantly fewer human-computer interactions in support of human target recognition (HTR) and significantly more communications and HCI interactions in support of Strike. Observations by the first author included repeated statements by the novices about the video game nature of the experimental runs and the C^2 prototype, and that they planned to approach each run from this "video game" perspective. Concerns about lost lives and equipment caused by the lack of proper identification are perhaps best sharpened by the actual combat experience of the experts.

Technical Skills

The authors have claimed that the Cadet participants were also technical novices on the C^2 prototype. The claim was supported by the investigation of differences between novice and expert performance times for long duration HCI tasks. The novice participants may have had more general computer and gaming experience than the expert participants, however, their lack of technical experience with the C^2 prototype may have resulted in slower performance on key tasks. As reported, the novices performed interactions in support of creating ground routes

significantly more slowly than the experts, and their performance times were significantly more varied in performing HCI tasks to create routes for air sensors. It appears that the novice command group's general computer technical skills could not substitute for the formal training and increased experience the experts had received on the C^2 system.

The task workload ratings from the Human Systems Integration Questionnaire may further explain performance differences between the two command groups. As reported, the novices rated workload higher on every task examined in the questionnaire, and in fact, 10 of the features received workload ratings from the novices that were twice as high as the rating provided by experts. This seems an interesting pattern, given the relatively computer-savvy stereotype of younger users, and the levels of computer and gaming experience reported by the novices. The novices' higher workload ratings may also serve as an indicator that general computer experience cannot be a substitute for requisite technical skills associated with the C^2 prototype. The C^2 prototype specific training is necessary to create technically proficient command groups.

Cognitive Skills

In the introduction, reference was made to the notion of schemas, in which expert knowledge is organized in such a way to enable more proficient and qualitatively different performance when compared to novice performance. If the development of tactical thinking skills follows a consistent and discernable pattern, then individual performance levels can be diagnosed, and training can be more efficiently targeted to individual needs. One consistent finding (Deckert, Entin, Entin, MacMillan & Serfaty, 1994; Ross, Battaglia, Phillips, Domeshek & Lussier, 2003) is that novices tend to focus much greater attention on friendly forces than on the enemy. In the research by Deckert, et al. just cited, officers judged as poor tacticians were described as all but ignoring the enemy. Better tactical performers showed a more balanced consideration of friendly and enemy forces. With superior tacticians the finding may reverse with greater consideration of unpredictable enemy actions compared to more predictable friendly actions. That finding was supported in this research where the novices communicated significantly more about friendly assets and significantly less about enemy assets than the experts.

It was also found that the novice command group in this research was focused primarily on a single dimension of enemy (location) compared to the experts who concerned themselves with a greater number of enemy characteristics. The finding has been noted in other fields, e.g., chess, where it is a frequent observation that novices focus on their own plans and moves and seem to ignore what the opponent is doing. One explanation, based on cognitive organization (i.e., schemas) is that in order to act one must consider 'own forces,' and such consideration virtually exhausts the capacity of the novices to build, maintain, and operate their mental models. Only with increasing expertise are models of sufficient complexity to encompass both 'own forces' and 'enemy forces' possible. Another explanation (Ross, et al., 2003) is that novices in all domains of expertise have a tendency to jump to solutions before gaining a sufficiently deep understanding of the situation.

An example of another type of difference in novice and expert performance is furnished by use of the Intervisibility Plotting (IV) tool, a feature of the FCS C^2 prototype used to determine line-of-sight between entities or areas on the battlefield. Correct use of the IV tool should directly support the ability of novices in particular to analyze terrain issues including observation, cover, and approach avenues. The HCI analysis confirmed that experts in the *planning* phase frequently used the IV tool to determine optimal placement of friendly vehicles. The IV tool was rarely used however during the *execution* stage by either experts or novices. Yet, the expert participants often discussed line-of-sight issues with respect to enemy and friendly vehicle locations. In particular, the experts were repeatedly concerned whether or not an enemy vehicle was within line-of-sight of a friendly platform. Notably, the IV tool is an example of how a technical capability can supplement tactical skills. At some point in the development of expertise, a tool that greatly supports novices may no longer be used by experts, at least when time is limited. Given the potential for shortcomings in the terrain data base or system crashes, the ability to "read" terrain to determine line-of-sight remains a skill required by experts.

System Design Recommendations

As reported, the novices and experts were asked to provide recommendations to the C^2 prototype's default settings on the Human Systems Integration Questionnaire. The novices provided many useful re-design recommendations for the C^2 prototype (see Appendix H). Input provided by lower ranking military personnel, like the Cadets, provides system designers with the information necessary to adapt the system to be more user friendly to individuals at entry technical and tactical skill levels. The novice command group was more representative of the type of personnel envisioned to command this future force, and their recommendations on ergonomic and software design issues could be used to create an interface more suitable for the target audience.

For example, the C^2 prototype is currently able to provide users a multitude of participant created alerts. Most alerts must be activated by each individual participant, however, and the activation of too many features can overwhelm participants with alerts. Current alert functions should be adapted to avoid alert overload and ensure more informative feedback when alerts are activated. For example, if a trainee or participant consistently fails to adjust the search radius for LAMs, the system could produce an alert immediately after the user has pushed the fire button requesting confirmation of the search radius entered. In addition, an automated mentor or wizard could be developed to answer technical and tactical questions during training, as well as during mission execution.

Training

As reported, the experts were provided three days of formal technical and collective training with the C^2 prototype, and the novices' technical experience varied from participating in demonstrations and pilot tests, to no technical training at all. By definition, experts are practitioners who have mastered a domain, occupation, or skill. Explanations of expertise stress that mastery of basic skills or "fundamentals" is prerequisite. For example, research on automaticity indicates that the performance of basic skills by novices versus experts is

distinguished by their respective reliance on controlled versus automatic processing (Fisk & Rogers, 1992). A growing consensus about experts is that their expertise entails automated basic skills, such as pattern recognition, that support more controlled skill-based solutions and decisions, such as situation recognition.

Training to mastery on basic skills is not complete until tactical and technical skills are well integrated. Training directed at the integration of basic skills should result in the expert-like automated processing of lower-level skills (Fisk & Rogers, 1992). Experts are then free to direct their limited controlled processing capabilities to the higher-order skills required of command groups, including visualizing the current and future battlefield situations. When basic skills are not well integrated performance decrements generally occur when operational requirements are time compressed (Morrison, Kelly, Moore & Hutchins, 1998) or made more complex as in the Too High experimental runs for FCS C^2.

Both command groups reported that more training was required. Inadequate individual and particularly collective training may have undermined the understanding and employment of the C^2 system's capabilities. Also, both command groups suggested that an increase in operational collective training would foster improved group cohesion, and therefore, increase group communication and performance. Inadequate collective training may have contributed to the novices' more decentralized and autonomous command structure.

Particularly at the small unit level, the transformation stressed by FCS requires that a relatively small command group must be able to command and control a complex mix of manned and autonomous systems. This is a daunting challenge for novices, and even for experts. Going beyond the current findings, after Experiment 4 when the expert participants had completed 40+ experimental runs, they expressed strong concerns about training and their ability to fully exploit Unit Cell capabilities (Lickteig, Sanders, Lussier, & Sauer, 2003). Whether for novice or expert, training must keep pace with the inevitable upgrades in software that revise or "drop" tools and features in earlier versions of C^2 systems. During Experiment 3, for example, such changes resulted in repeated queries from the Battlespace Manager: "Where's the IV button?" The location of the IV tool/button had migrated in the upgrades between Experiment 2 and 3.

Part-task training often improves learning and performance of complicated tasks (Kirlik et al. 1998). ARI concludes that the complexity inherent in more advanced C^2 systems, such as the FCS C^2 prototype, represents a part-task training requirement for command group personnel. A part-task training capability, particularly for more complex and time consuming tasks, should help novices and experts learn and retain mastery on highly interdependent tactical and technical tasks.

Two part-task methods called segmentation and simplification are described below that have proven successful in training more complicated C^2 system skills (Kirlik et al., 1998). For each method, training examples based on the FCS C^2 prototype are provided to more clearly specify the training requirement.

The first part-task training method, segmentation, involves reducing a larger task into smaller parts, namely individual steps and stages. In theory, temporal step-by-step segmentation

reinforces the outcome of the previous steps, and increases the likelihood that the trainee will understand the process needed for successful performance. This method also helps identify which partial tasks may require more training. A proposed example of how segmentation could apply to training on the FCS C^2 prototype follows.

The firing of a loitering attack missile (LAM) with the C^2 prototype requires four basic steps: proper identification of a suitable enemy target, opening the Netfire window, target acquisition, and selection and activation of special features associated with firing a LAM. Segmentation of each step required for successful deployment of a missile should benefit speed and accuracy of the whole task.

The first step of segmented training should address proper identification of a suitable target. For example, an enemy scout is not a suitable target for an attack missile, whereas a thick armored tank is a suitable target. Correct identification implies all HTR requirements have been met to ensure that a proper missile or other munition is allocated to the destruction of the enemy target. The second step of training should address the interactions needed for opening the Netfire window. HCI analysis through Experiment 4 identified three ways to achieve this, which are: a toolbar button for the Netfires, a "Quick Fire" toolbar button that controls all munitions available to the participants, and a menu option made available by right clicking the Netfire platform icon. The third step of training should address proper target acquisition. The HCI analysis identified three types of target acquisition, which are: selection of a target by clicking the platform icon, selection of the target by selecting the target's system or user provided identification tag, and the selection of a target by a drop down menu provided within the Netfire window. The fourth step needs to address the special features provided by the system when firing a LAM. These features include: setting the search radius, programming the LAM to loiter or attack, and setting various time or platform conditions to the execution of the fire.

The second part-task training method, simplification, entails temporary removal of selected elements in a complicated or time consuming task that could interfere with performance or confuse training objectives (Kirlik et al., 1998). As mastery on more simplified task elements is achieved, additional elements should be added until the trainee is fully and effectively trained on all aspects of the whole task. Simplification is most useful for tasks that are initially difficult to perform, or high workload tasks. Simplification not only improves whole task performance, but is often a more efficient and less frustrating method for achieving skill mastery (Kirlik et al. 1998).

One example of simplification would be the temporary reduction of the overall task requirements for a complicated duty position, such as the Information Manager. For FCS C^2, the Information Manager is the primary source for all reconnaissance and intelligence information provided to the Unit Cell command group. This job includes performing at least three separate task areas, including: movement of aerial sensors, tasking aerial sensors for reconnaissance, and performing image analysis in support of HTR and BDA. During preliminary training exercises with the C^2 prototype, the Information Manager trainee could be required to perform only one of the three listed task areas, while the other tasks are performed by expert trainers or an automated system. When the selected task area is mastered, training should shift to either of the two remaining task areas, and so on until the trainee is successfully performing all duty position

requirements. A related example would be training that allows the Information Manager to master control over one of four Micro UAVs before presenting the requirement to control multiple Micro UAVs.

Feedback

Feedback is a multi-purpose training requirement that: informs the trainee and trainer about skill levels and deficiencies, provides valuable information on mistakes and alternatives, reinforces training goals, and promotes and guides mastery.

For the FCS C^2 experiments, feedback for participants was primarily provided by human trainers, expert technicians, and subject matter-experts in AARs. Although this feedback was beneficial, extensive improvement in the nature and quality of feedback is required for more effective training. Trainers tend to base their feedback on overall performance instead of the combination of preceding factors that led to the performance (Kirlik et al., 1998). Preceding factors are usually key performance limitations, such as misallocation of priorities, or specific events or interactions that created the performance limitation. Ideally, feedback should help trainees identify the cause or source of errors, not just the error "symptom" or the consequences of the error. However, feedback that effectively diagnoses the source or cause of errors often requires a very high level of attention to the pattern and details of performance. Such attentiveness severely taxes human instructors, particularly with higher trainee-to-instructor ratios. For example, novice participants during the Summer Experiment received many warnings by Blue Higher on their failure to properly identify targets before engaging. However, the more general underlying problem seemed to be the novices' failure to maintain situational awareness on all battlefield entities, not just the enemy. Recall, the novices not only performed less HTR in support of target identification but also fewer target, friendly, and area query interactions to support their general situational awareness.

Conclusion

The preliminary findings on novice and expert performance reinforce the need for a transformation in Army training, particularly to develop the demanding tactical and technical skills required in small command groups of the future. The results provide an emerging empirical database to help identify FCS command group tasks and training requirements. Overall, the research findings identify key differences that will need to be understood more completely in order to design future training capable of meeting the new training requirements for FCS and transforming novices into experts in future command groups.

References

Borko, H., & Livingston, C. (1989). Expert-novice differences in teaching: Cognitive analysis and implications for teacher education. *Journal of Teacher Education*, 40, 36-42.

Campbell, C. H., Quinkert, K. A., & Burnside B. L. (2000). Training for performance: The structured training approach. (ARI Special Report 45.) Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Cordesman, A. H. & Wagner, A. R. (1996). *The lessons of modern war, Volume IV: The gulf war.* Boulder, Co: Westview Press.

Deckert, J. C., Entin, E. B., Entin, E. E., MacMillan, J., & Serfaty, D. (1994). *Military decision-making expertise: Final report.* Fort Leavenworth, KS: U.S. Army Research Institute.

Defense Advanced Research Projects Agency. (2001). Future combat systems: CSE functions manual (draft). FCS Unit Cell C^2 Study Technical Team, Fort Monmouth, N.J.

Durbin, D. B. (2002). Assessment of the human factors characteristics of the AH-64D Apache Longbow crew stations. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Ericsson, K. A., & Simon, H. A. (1984, 1993). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Fisk, A. D., Kirlik, A., Walker, N., & Hodge, K. (1992). *Training for decision making under stress: Review of relevant literature and research plan* (HAPL-9201). Atlanta, GA: Georgia Institute of Technology, School of Psychology, Human Attention and Performance Laboratory.

Fisk, A. D., & Rogers, W. A. (1992). The application of consistency principles for the assessment of skill development. In J. W. Regian & V. J. Shute (Eds.), *Cognitive approaches to automated instruction* (pp 171-194). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hershey, D. A., Walsh, D. A., Read, S. J., & Chulef, A. S. (1990). The effects of expertise on financial problem solving: Evidence for goal directed, problem solving scripts. *Organizational Behavior and Human Decision Processes*, 46, 77-101.

Hornik, S. & Ruf, B. M. (1997). expert systems usage and knowledge acquisition: An empirical assessment of analogical reasoning ion the evaluation of internal controls. *Journal of Information Systems*, 11(2), 57-74.

Kirlik, A., Fisk, A. D., Walker, N., & Rothrock, L. (1998). Feedback augmentation and part-task practice in training dynamic decision making skills. In *Making Decisions Under Stress: Implications for Individual and Team Training*, edited by J. A. Cannon-Bowers, and E. Salas. American Psychological Association: Washington, DC, 91-113.

Kirlik, A., Walker, N., Fisk, A. D., & Nagel, K. (1996). Supporting perception in the service of dynamic decision making. *Human Factors*, 38(2), 288-299.

Kramer, S. H. (1999). When are two heads better than one? The role of expertise and task difficulty in individuals, statistical group, and interacting group problem solving. *Dissertation Abstracts International: Section B: The Sciences and Engineering*. 60(3-B), 1350.

Lickteig, C. W., Sanders, W. R., Durlach, P. J., Rainey, S. J, & Carnahan, T. J. (2002). Future combat systems command and control (FCS C^2) human functions assessment: Experiment 2 interim report. Available from Program Manager (PM) FCS C^2.

Lickteig, C. W., Sanders, W. R., Durlach, P. J., & Carnahan, T. J. (in preparation). *Future combat systems command and control (FCS C^2) human functions assessment: Experiment 3 interim report.* (ARI Research Report). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Lickteig, C. W., Sanders, W. R., Lussier, J. W. & Sauer, G. (2003). *A Focus on Battle Command: Future Command Systems Human Systems Integration.*. Paper HS217 submitted for the 2003 Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL.

McKinney, E. H. (1993). Flight leads and crisis decision making. *Aviation, Space, and Environmental Medicine*, 64, 359-362.

Morrison, J. G., Kelly, R. T., Moore, R. A., & Hutchins, S. G. (1998). Implications of decision-making research for decision support and displays. In *Making Decisions Under Stress: Implications for Individual and Team Training*, edited by J. A. Cannon-Bowers, and E. Salas. American Psychological Association: Washington, DC, 91-113.

NASA-Ames Research Center, Human Performance Group. (1986). Collecting NASA workload ratings: A paper and pencil package (Version 2.1). NASA-Ames Research Center. Moffet Field, California.

Pronti, J., Molnar, S., Wilson, D. (2002). C^2 architecture study: Interim report. Available from Program Manager (PM) FCS C^2.

Ross, K. G., Battaglia, D., Phillips, J., Domeshek, E., & Lussier, J. W. (2003). Mental Models Underlying Tactical Thinking Skills. Paper submitted for the 2003 Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL.

Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. *Human Factors*, 27(3), 285-300.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, 8(4), 384-395.

Vosniadou, S. & Ortony, A. (1989). Similarity and analogical reasoning: A synthesis. In *Similarity and Analogical Reasoning*, edited by S. Vosniadou, and A. Ortony. Cambridge, U.K.: Cambridge University Press, 1-15.

Wiggins, M. & Henley, I. (1997). A computer-based analysis of expert and novice flight instructor preflight decision making. *The International Journal of Aviation Psychology*, 7(2), 365-379.

Appendix A

List of Acronyms

A-160	Unmanned Rotary Wing Observation Platform
AAR	After Action Review
AFRU	Armored Forces Research Unit
AGM	Attack Guidance Matrix
AH-64D	Apache Longbow Helicopter
ARI-Knox	U.S. Army Research Institute for the Behavioral and Social Sciences at Fort Knox
ARL	Army Research Laboratory
ATR	Aided Target Recognition
BCV	Battle Command Visualization
BDA	Battle Damage Assessment
BLOS	Beyond Line-of-Sight
C^2	Command and Control
C^2VEH	C^2 Vehicle
C4ISR	Command, Control, Communications, Computer, Intelligence, Surveillance, and Reconnaissance
CECOM	U.S. Army Communications Electronics Command
CIC	Combat Information Center
CM	Collection Management
COA	Course of Action
COFT	Conduct of Fire Trainer
CSE	
DARPA	Defense Advanced Research Projects Agency
FBC	Future Battlefield Conditions
FCS	Future Combat Systems
FCS C^2	Future Combat Systems Command and Control
FW	Future Warriors
GCM	Graphic Control Measure
HCI	Human-Computer Interaction
HSIQ	Human Systems Integration Questionnaire
HTR	Human Target Recognition
IV	Intervisibility Plotting
LAM	Loitering Attack Missile
LOS	Line-of-Sight

LTC	Lieutenant Colonel
METT-TC	Mission, Enemy, Troops, Terrain, Time, Civilians
NAI	Named Area of Interest
NASA	National Aeronautics and Space Administration
NLOS	Non-Line-of-Sight
NTC	National Training Center
OneSAF OTB	One Semi-Automated Forces Testbed Baseline
PAM	Precision Attack Missile
PM	Program Manager
RDEC	Research and Development Center
ROTC	Reserve Officer Training Candidate
SD	Standard Deviation
SME	Subject Matter Expert
STO	Science and Technology Objective
TLX	Task Load Index
UAV	Unmanned Aerial Vehicle
UGS	Unmanned Ground Sensor

Appendix B

Verbal Communication Coding Scheme: Definitions and Examples

For each chunk select

SOURCE (for each verbal chunk select one and only one Source code)

1	Within Cell (Black)	Cell = 4 C^2 prototype operators.
2	Cell <-> Blue (Team)	
3	Cell <-> White (Higher)	
4	Cell<->Subordinate	Subordinate (includes C^2Vehicle gunner & driver).
5	Blue<-> White	
6	More than 2-way (e.g., Cell<->White<->Blue)	Only to be used in cases where more than 2 elements involved in SAME conversation.
7	Other	E.g., to technical support people.

FUNCTION (for each verbal chunk select one and only one Function code)

1	See	Detect or identify enemy or friendly positions, or significant terrain aspects (not BDA).
2	Plan	Interpret data, predict enemy COA, generate own COA
3	Move	Manage/monitor/control asset movement.
4	Strike	Manage/monitor/control lethal/nonlethal effects.
5	BDA	See for purposes of BDA.
6	Other	None of the above.

VALENCE (for each verbal chuck select one and only one Valence code)

	1	0	-1
See	Ability to See	Neutral/inconclusive	Inability to See
Plan	Plan Working	Neutral/inconclusive	Plan not Working
Move	Ability to Move	Neutral/inconclusive	Inability to Move
Strike	Ability to Strike	Neutral/inconclusive	Inability to Strike
BDA	Ability to Confirm Kill	Neutral/inconclusive	Inability to Confirm Kill
Other	Other Function Achieved	Neutral/inconclusive	Other Function not Allowed

TYPE (for each verbal chunk select one and only one Type code)

1	Share. Verbalization about what is seen or known.
2	Action. Verbalization about what speaker is doing at the moment—verbalization accompanying action such as fire or move.)Not the decision process. Not actions as I see, monitor, track, etc. Not describing someone else's actions.)
3	Direction. Order, command, delegation of responsibility.
4	Ask. Verbalization begins with request for information, confirmation, assistance, or assets and ends with either informational answer or no response, with little or no discussion. Not rhetorical questions.
5	Process. Infer, synthesize, fuse, understand, turn data into information without consequent decision or direction. Can start with Share, Action, or Ask.
6	Decide. Like Process, but in addition, includes a verbalized decision or plan.
7	Other.

Verbal Communication Coding Scheme: Definitions and Examples
(continued)

FACTOR (for each verbal chunk select one and only one Factor code)

MISSION	
1	Original Plan: Concerning mission goals and plans prior to execute phase.
2	Dynamic Planning: Tactical re-planning during the execute phase in response to changing events and available assets. Must have stated COA (course of action). Changes from Original Plan.
3	Situational Understanding. Integration/summary of current situation involving multiple factors; but without stated COA.
ENEMY:	
4	Location: Sensor hit(s) – locate enemy positions.
5	Identification: Identify targets – identify nature of enemy target.
6	Disposition: Probable enemy COA, strategy, or tactics.
7	BDA: Battle Damage Assessment – cell seeks/discusses feedback on damage they inflict on enemy.
TERRAIN	
8	When terrain is the prime focus (e.g., can we travel over that kind of terrain?, we should go this way because it will provide cover). Example: "Moving to low ground." Not simply map locations (e.g., not, sensor hit north of the wall).
TROOPS and Assets (Soldiers, Equipment, Vehicles) Friendly only.	
9	Location Status: Position report/assessment.
10	Movement Status: Mobility report/assessment (includes fuel).
11	See Status: Sensor report/assessment.
12	Strike Status: Fire power report/assessment (includes # of remaining missiles).
13	Communications/network functionality (radio, internet, or other; cell to outside cell, including semi-autonomous sensors).
14	Information management systems: C^2 prototype user interface tools.
15	Survivability Concern: Asset in danger.
16	Survivability Move: Defensive move to remove asset from immediate danger.
17	Loss/Casualty: Asset destroyed (catastrophic hit).
18	Move Action: Move/Manage/Maneuver [Active, Not position report] Excluding Survivability Move; Also See Terrain.
19	Strike Action Lethal: Launch/fire/deploy with intent to destroy (includes LAMs)
20	Strike Action Nonlethal: Launch/fire/deploy (could include unarmed sensors, propaganda, smoke, jamming of enemy, etc.).
21	Training (Soldier training, mission rehearsal).
22	Other-- having to do with troops or assets but none of the above.
TIME	
23	When time is the prime focus (e.g., how much time something will take, how much time is available, order of priority, synchronization of actions).
CIVILIANS	
24	Any issues regarding how to deal with civilians: avoiding, provisioning, protecting, etc. Not mere sensor hits of civilians, unless first time mentioned.
Other	
25	Other (e.g., humor, personal, leadership, morale).

Verbal Communication Coding Scheme: Definitions and Examples (continued)

Coding rules of thumb for Verbal Communication Rating Scheme:

Type:

Rationale: Share, Action, and Direction are meant to be relatively short interactions, without a lot of discussion. Chunks including a lot of discussion or consideration of multiple aspects of situation should be either Process or Decide. These are distinguished by whether there is a definite conclusion reached (Decide) or not (Process).

1. When in doubt between Share and Action, choose Share.
2. When in doubt between Share and Process or Decide choose Process or Decide (as appropriate).
3. When in doubt between Ask and Process or Decide choose Process or Decide (as appropriate).
4. Rhetorical questions or questions following an announcement should **not** be coded as Ask. (e.g., I have a mover, do you see it?—should be Share).
5. You have to pay attention to both the beginning and end of a chunk. If it contains a verbalized decision, plan, or direction, it is Direction or Decide, regardless of how it begins. Distinguish Direction and Decide by whether it is preceded by some discussion (Direction is not; Direction stands alone. Decision is preceded by relevant discussion). For example, a sensor hit followed by a direction to fire should be classed as a (type: direction/subject: lethal effects, not type: share/subject: sensor hit. (as per example below).

- We've identified the other mover wheel coming out of hidden valley is a URAL (truck).
- Dave take that URL with a PAM.

Subject:

Rationale: Choose the major subject of the chunk. Consider, what information is the speaker trying to convey?

1. Choose Dynamic Replanning or Situation Understanding when the conversation contains discussion of multiple assets. Use Dynamic Replanning when it does include a course of action (here's what we should do). Use Situation Understanding when it does not include a course of action or plan, but only summarized the current situation. If chunk contains discussion of only a single asset, choose the appropriate category related to that type of asset or action.
2. Sometimes judgment will be required. In these cases try to imagine what is the subject the speaker is trying to convey? (e.g., "Darya found by Roboscout" Context will usually help. I would tend to code this as Sensor hit, especially if it is the first time the hit was mentioned. On the other hand, a preceding question regarding which sensor system detected the Darya, would make the main information conveyed by the utterance Sensors…" Who found the Darya? Darya found by Roboscout").
3. When the Type is scored as "Ask,", and there is ambiguity as to Subject, focus on what kind of information the asker is after.

System:

1. If more than one asset is mentioned, give more than one system code. Note the existence of categories such as Other, Unspecified, or Not Applicable.

- Other—asset is mentioned, but is not one of the choices.
- Unspecified—clearly talking about one of the assets but you can't tell which (e.g., you know it is a lethal effect, but you don't know which one).

Not Applicable—a system is simply not applicable to the subject discussed in the chunk.

Appendix C

Examples of Verbal Chunks from Experiment 3 by Function with assigned Valence Values

Function Valence	
Plan	
1	I was just thinking about the birds being too far up there, up North. Bring them South then. I will, I don't control anything, I've got to ask team to bring them further South. And I can do that, sir, if you don't mind. No go right ahead, keep them South, that's fine with me.
0	I got an idea, do you want to try something new? No.
-1	We have unconfirmed as of yet BDA on a tank in the South and a couple of tanks in center sector but we don't have enough intelligence yet to give us a good read of the battlefield other than the fact that he tried to move forward in the center, and I will keep you informed.
See	
1	There's an unknown radio hit.
0	They haven't fired any artillery yet, have they? (no response)
-1	Dang, there's nothing in my images here.
Move	
1	So we need to get those 2 micros back down there. They're coming down.
0	Where is the SAR bird? (no response)
-1	That one's stuck there, number 2 is just not… microUAV 2 is not responding.
Strike	
1	Did you? Did you fire 4? Yeah, I just fired 4.
0	Well, the question is, do we reengage? (no response)
-1	OK. Interestingly, you lost comms on the PAMs that you sent. You see that? PAM 54 lost comms. It didn't attack. Hold on a second, the last 2 PAMs that you sent lost comms and did not go to the target. You want me to show you? The one on the Darya did that too. Neither one hit anything. They both lost comms.
BDA	
1	Here is a better image. It looks like it might be perhaps a fire power kill, maybe a fire power, mobility kill.
0	PAM 16, where did that hit? (no response)
-1	Is it broke? Did we kill it?' I don't know, it doesn't look like it's broke from this image right here, it's hard to tell.
Other	
1	What's the red dot mean? It means that's where it detected something and takes a picture, or that's the place where the Garm was templated.
0	Blue 6, Black 6 (no response)
-1	I've got a right screen frozen.

Appendix D

Human-Computer Interaction Coding Scheme

100 PLAN

110. Create/Update a Mission and COA
    111. Create Overlay Graphics and Map Annotations
    112. Place platforms on the map
    113. Rehearse the Plan
    114. Execute the Plan
    115. Point on Map Using Cursor/Indicate an Area
    116. Move icons on map using drag/drop
    117. Modify overlay graphics

120. Alerts
    121. Create Alerts

200 MOVE

210. Move Ground Assets
    211. Create routes
    212. Start, Halt or Resume a platform
    213. Edit an existing route
    214. Delete all tasks
    215. Place Unmanned Ground Sensor (UGS)
    216. Overwatch
    217. Generate Route
    218. Recon an Area
    219. Dismount Future Warrior Team

220. Move Air Assets
    221. Create Routes
    222. Delete all tasks
    223. Edit an existing route
    224. Recon an Area/Auto Recon Targets
    225. Task to Hover

230. Group Follow
    231. Create Ground Follow
    232. Create Air Follow
    233. Create Mixed (Ground and Air) Follow

300 SEE

310. Manipulate Map
    311. Zoom Map
    312. Scroll Map
    313. Save Map Zone (Preset Zoom)
    314. Select Preset Zoom

320. Use Visualization Aids
    321. Toggle Range Fans
    322. Plot Intervisibility
    323. Measure Distance
    324. Display on Heads Up
    325. Select/Change Windows or window area for display
    326. Change Graphic Control Measure(GCM) Settings
    327. Move Visual Reference Points
    328. Toggle Polygon Layer
    329. Toggle Polyline Layer

330. Display Sensor Data
    331. Display Target Catalog
    332. Query Enemy
    333. Query Friendly
    334. Query Area
    335. Change Sensor
    336. Toggle sensor fans
    337. Acknowledge Alert Window
    338. Highlight Target on Map using Target Catalog
    339. Take Manual UAV Picture

340. Recognize Targets
    341. Display target images
    342. Refine image
    343. Change Map Icons to reflect target status
    344. Change State View
    345. Select recon target by clicking icon
    346. Select recon target by select window
    347. Display recon location image

350. Assess Battle Damage
    351. Display target images
    352. Refine image
    353. Change Map Icons to reflect status
    354. Assign by BDA Recommendations
    355. Acknowledge Alert

360. Summarize Situational Awareness
    361. Open Threat Management Window
    362. Query Enemy (Unit Viewer)
    363. Track Unit (Unit Viewer)
    364. Query Friendly (Unit Viewer)
    365. Update/Change information displayed in Battlefield Assistant
    366. Acknowledge Instant Message
    367. Type Instant Message
    368. Send Instant Message

400 STRIKE

410. New Features
- 411. Recommend Fire Unit
- 412. Open Quick Fire
- 413. Engage from Enemy Intel Window
- 414. Prohibit Attack

420. Target Designation
- 421. Designate by Icon Click
- 422. Designate by Menu Selection
- 423. Designate by "Select" Window

430. Fire Weapon System
- 431. Fire Netfire LAM
- 432. Reassign LAM
- 433. Fire Netfire PAM
- 434. Fire LOS
- 435. Fire C^2Vehicle (Gun and Javelin)
- 436. Fire FW CARRIER (IFV)
- 437. Fire Dismount Javelin
- 438. Delete all scheduled fire tasks

440. Monitor Fires Execution
- 441. Query LAM
- 442. Query PAM
- 443. Query LAM (Unit Viewer)
- 444. Query PAM (Unit Viewer)

450. Scheduled Fires
- 451. Set Minutes to Fire
- 452. Set Delimiters

460. Attack Guidance Matrix
- 461. Create Attack Guidance Matrix (AGM)
- 462. Select target category
- 463. Select Autofire threat level
- 464. Set Weapon Priorities
- 465. Edit Roles
- 466. Adjust threat index settings
- 477. Select AGM

500 OTHER MANUAL ACTS

510. General
- 511. Reboot system

Appendix E

Task Load Index Rating Scales

Task or Mission Segment: ____________________________________

Please rate the task or mission segment by putting a mark on each of the six scales at the point which matches your experience.

Mental Demand

Very Low — Very High

(HOW MENTALLY DEMANDING WAS THE TASK?)

Physical Demand

Very Low — Very High

(HOW PHYSICALLY DEMANDING WAS THE TASK?)

Temporal Demand

Very Low — Very High

(HOW HURRIED OR RUSHED WAS THE PACE OF THE TASK?)

Performance

Failure — Perfect

(HOW SUCCESSFUL WERE YOU IN ACCOMPLISHING WHAT YOU WERE ASKED TO DO?)

Effort

Very Low — Very High

(HOW HARD DID YOU HAVE TO WORK TO ACCOMPLISH YOUR LEVEL OF PERFORMANCE?)

Frustration

Very Low — Very High

(HOW DISCOURAGED, IRRITATED OR ANNOYED WERE YOU?)

AFTER RUN SURVEY

Part 2. CSE Effectiveness

How effective was the CSE in support of C^2 Functions and METT-TC Factors?

C^2 Functions	Very Ineffective	Ineffective	Borderline	Effective	Very Effective
1. PLAN: Create Mission/COA, tools (range fans, intervisibility, distance). Comments: ____	1	2	3	4	5
2. MOVE: Create routes, move, halt, retask ground and air assets. Comments: ____	1	2	3	4	5
3. SEE: Target data, Human Target Recognition, sensor data integration. Comments: ____	1	2	3	4	5
4. SHOOT: Plan and execute fires, check resources, BDA. Comments: ____	1	2	3	4	5

METT-TC Factors					
1. Mission: Operations Plan, dynamic planning, coordination w/ higher. Comments: ____	1	2	3	4	5
2. Enemy: Activities, composition, probable COA. Comments: ____	1	2	3	4	5
3. Terrain: Key Terrain, avenues, observation lines, fields of fire. Comments: ____	1	2	3	4	5
4. Troops And Assets: Training adequacy, vehicles, sensors, weapons. Comments: ____	1	2	3	4	5
5. Time: Maneuver, coordination, asset task time (ex. time of flight). Comments: ____	1	2	3	4	5
6. Civilians: Identifying, tracking, avoiding, provisioning, protection. Comments: ____	1	2	3	4	5

TRAINING ADEQUACY

Duty Position ________________ Date __________ Run # _________

How adequate was the training provided to prepare you **individually** for the C^2 cell, in terms of content coverage and time spent?

CONTENT: ____

____

____

TIME: ____

____

____

How adequate was the training provided to prepare the members of the C^2 cell **collectively** to work as a team, in terms of content coverage and time spent?

CONTENT: ____

____

____

TIME: ____

____

____

Human Systems Integration Questionnaire

Name ______________________ Duty Position ___________________ Date _________

This questionnaire asks for your feedback on how well the CSE/C^2 Prototype system supports human information requirements for battle command.

1. Using the workload description rating scale, provide a rating of your perceived workload by writing the corresponding number in the "Rating" column for each of the following *functions* (plan, move, see, strike) and related tasks from 1 to 10. Please add any other key tasks that impact workload by writing them in the cells labeled "Others?" below.

Workload

1	2	3	4	5	6	7	8	9	10
Insignificant	Very Low	Low	Task Attention Compromised	Reduced Attention	Little Spare Capacity	High	Very High	Extremely High	Task Abandoned

Function and Task	Rating (1-10)	Recommended Improvement to Reduce Workload
Plan		
Position Icons on the Map		
Rehearse the Plan		
Mission/Task Change		
Others?		
Move		
Ground Asset		
Air Asset		
Position Unmanned Ground Sensors (UGS)		
Task Sensor to Recon		
Others?		

WORKLOAD

1	2	3	4	5	6	7	8	9	10
Insignificant	Very Low	Low	Task Attention Compromised	Reduced Attention	Little Spare Capacity	High	Very High	Extremely High	Task Abandoned

Function/Task	Rating (1-10)	Recommended Improvement to Reduce Workload
See		
Map Manipulation (eg., Zoom)		
Using Heads Up Display		
Altering Workstation Layout		
Creating Alerts		
Collecting Sensor Data		
Interpreting Sensor Data		
Collecting Picture Data		
Interpreting Picture Data		
Others?		
Strike		
Detect Targets		
Recognize Targets		
Classify Targets		
Identify Targets		
Designate Target		
Fire NETFIRES		
Fire Line-of-Sight (LOS)		
Battle Damage Assessment (BDA)		
Others?		

2. Please check Yes or No for each item in column 2a (task clarity) and 2b (task complication) below.

a. Are there any tasks that are not logical or consistent?	b. Did any tasks require too many steps to complete?
Share information on Heads Up Display Yes___ No___	Share information on Heads Up Display Yes___ No___
Create routes Yes___ No___	Create routes Yes___ No___
Edit existing tasks Yes___ No___	Edit existing tasks Yes___ No___
Measure distance Yes___ No___	Measure distance Yes___ No___
Tailor workstation (window size and log out) Yes___ No___	Tailor workstation (window size and log out) Yes___ No___
Changing sensor used Yes___ No___	Changing sensor used Yes___ No___
Human Target Recognition (HTR) Yes___ No___	Human Target Recognition (HTR) Yes___ No___
Target designation Yes___ No___	Target designation Yes___ No___
Allocating search radius (LAM/PAM) Yes___ No___	Allocating search radius (LAM/PAM) Yes___ No___
Fire NETFIRES Yes___ No___	Fire NETFIRES Yes___ No___
Fire LOS Yes___ No___	Fire LOS Yes___ No___
Fire Javelin Yes___ No___	Fire Javelin Yes___ No___
Battle Damage Assessment (BDA) Yes___ No___	Battle Damage Assessment (BDA) Yes___ No___

3. Please place a checkmark in the appropriate box that corresponds to the amount of trouble you experienced viewing the following display characteristics.

Display Characteristics	Never Have Trouble	Seldom Have Trouble	Occasionally Have Trouble	Frequently Have Trouble
Legibility of text				
Contrast between symbols and background				
Brightness of displays				
Size of displays				
Color of symbols				
Text on displays				

4. Please place a checkmark in the appropriate box that corresponds to the amount of difficulty/ease in accomplishing the following tasks.

Symbology Characteristics	Very Easy	Somewhat Easy	Borderline	Somewhat Difficult	Very Difficult
Ease of distinguishing between friendly and threat icons					
Ease of distinguishing between moving and stationary threat icons					
Ease of visualizing past and future threat positions					
Ease of distinguishing between LAM/PAM missile icons					
Ease of visualizing missile trajectory and intended target					
Ease of determining what entity detected the threat target					
Ease of understanding navigation symbology (waypoints, hazards, etc.)					

5. Are there any menus or tasks that are either too difficult to navigate or take too much time that they complicate your performance? Please circle: **YES** **NO**

Please explain:

____

____

____

____

____

6. Are there any items in the CSE/C^2 Prototype improperly or confusingly labeled? Please circle: **YES** **NO**

Please explain:

____

____

____

____

____

7. Please place a checkmark in the appropriate box that corresponds to how easy is it to read/understand the data provided in the following CSE/C^2 Prototype windows.

	Very Easy	Easy	Borderline	Difficult	Very Difficult
Map Window					
Mission Workspace (Table of Organization and Equipment)					
Execution Window (mission timeline)					
Resource Availability					
Asset Window					
Alert Ticker Window					
Alert Editor					
Target Catalog					
Battlefield Assistant					
Graphic Control Measures (GCM) Window					
Collection Management (CM) Planner					

8. Are there any CSE/C^2 Prototype default settings that you would recommend a change? If there is no change that you recommend for a given default, please place a check mark in the "No Change" column for that default. The following blank cells are for your suggestions on other defaults that need adjustments.

Default	No Change	Recommended Change
Search Radius for LAM/PAM		
Should out of range ammo be an option		
When out of a type of ammo, should it continue to be the default or even an option		
Should friendly entities be a default for weapons (RoboScout, UGS)		
Should picture adjustments be saved for later viewing by you or a different person		
Should your tailored settings and user preferences be saved and automatically loaded when the system crashes		
Graphics such as the Named Area of Interests (NAIs) and phaselines being click/drag active during execution phase		
Others?		

Appendix F

Novice Duty Position Responsibilities

	Cell Commander	North Ground Commander	South Ground Commander	Information Manager
Battlefield Area	Entire Area	Northern Area	Southern Area	Entire Area
Functions	Make sure all tasks are being accomplished	Direct engagement of all ground forces north of mid-field	Direct engagement of all ground forces south of mid-field	HTR and BDA of Battlefield
C^2 Cell Assets				
Shadow	Task for Recon			
Command and Control Vehicle	Control Movement			
MicroUAV's (4)	Controlled 2 for Recon			Controlled 2 for Recon
Line-of-Sight Vehicle (2)		Controlled all movements, tactically and strategically	Used to destroy enemy	
Netfires Platform (2) (LAM, PAM, UGS)		Still, long range firing system	Destroy all vehicles in my engagement area	
Roboscout Vehicle (2)		Used as short range scouts	Moved forward of units to detect enemy	
Future Warrior Vehicle (2)		Used as short range scouts and Javelin offensive weapons	Used as a sensor and weapon, but generally assessed no specific tasks	
Future Warrior Team (4)		Used as short range scouts and Javelin offense weapon	Utilize as scouts out in from to ID enemy	
Javelin Team (2)		Offensive weapon		

Appendix G

Expert Duty Position Responsibilities

	Cell Commander	Battlespace Manager	Information Manager	Effects Manager
Battlefield Area	Far Fight	Close Fight	Close Fight	Close/Middle Fight
Functions	Seeing, understanding the battlefield. Developing, preparing, synchronizing COAs, using C^2 to execute plans.	Primary - Movement of both Roboscouts; LOS; and FIFVs with associated Infantry. Alternate – NETFIRES and Command.	Primary - Intelligence, identify enemy COA. Alternate – Can move vehicles.	Primary – Plan and execute NETFIRES fires, UGS placement, monitor ammo expended. Alternate – Plan and execute UAV movement.
C^2 Cell Assets				
Shadow			Primary operator, includes DVO imagery analysis	
Command and Control Vehicle	Directs movement			
MicroUAV's (4)			Controls all 4 UAVs, primary imagery analysis	Imagery analysis
Line-of-Sight Vehicle (2)		Does not have time for imagery analysis	Imagery analysis	Imagery analysis
Netfires Platform (2) (LAM, PAM, UGS)	Plans UGS placement		Execute UGS	Controls all 60 NETFIRES rounds
Roboscout Vehicle (2)		Direct, manage, and monitor	Imagery analysis	Imagery analysis
Future Warrior Vehicle (2)		Direct, manage, and monitor		
Future Warrior Team (4)		Direct, manage, and monitor		
Javelin Team (2)		Direct, manage, and monitor		

Appendix H

Novice Default Setting Recommendations

Default Duty Position	Recommended Change
Search Radius (LAM/PAM)	
Commander	Be able to change default radius (depending on battle conditions)
North Manager	800 meters to one click.
Should out of range ammunition be an option?	
Information Manager	No, it makes no sense for it to be an option.
When out of a type of ammunition, should it continue to be the default or even an option?	
Commander	This should definitely not be an option.
North Manager	It shouldn't be an option.
South Manager	No, it should be removed.
Information Manager	No, remove it from the list.
Should friendly entities be a default for weapons?	
Commander	No. Remove them from the list.
North Manager	It shouldn't be an option
South Manager	No, they shouldn't be an option.
Information Manager	No.
Should picture adjustments be saved for later viewing by you or a different person?	
Commander	Yes, pictures should be saved.
South Manager	Yes, they should be saved.
Information Manager	Yes.
Should your tailored workstation be saved when the system crashes?	
Commander	Yes.
South Manager	Yes.
Information Manager	YES!
Should graphics such as the NAIs and phaselines be click/drag active during execution phase?	
Commander	Only be able to right click on them to for editorial purposes.
South Manager	No, they shouldn't be active at all.
Information Manager	No
Others	
North Manager: *Need a set map size that you can click a button and go back to "Normal" from a zoom in state.* *Need to know whether I looked at an image or if someone else has looked at it already.* *Have a button to toggle off picture icons on the main display to remove screen clutter.*	
Information Manager: *Be able to task assets from Resource Availability window.*	

Note: If a command group participant is omitted under a question, he recommended no change to the default setting.

Appendix I

Expert Default Setting Recommendations

Default Duty Position	Recommended Change
Search Radius (LAM/PAM)	
No Change	
Should out of range ammunition be an option?	
No Change	
When out of a type of ammunition, should it continue to be the default or even an option?	
Effects Manager	Would be nice to see more than one recommended fire option.
Should friendly entities be a default for weapons?	
Battlespace Manager	Should not be an option in any targeting cue.
Commander	(No written recommendation for improvement)
Should picture adjustments be saved for later viewing by you or a different person?	
Commander	Yes
Battlespace Manager	Different person.
Information Manager	(No written recommendation for improvement)
Should your tailored workstation be saved when the system crashes?	
Commander	Yes
Battlespace Manager	Yes
Information Manager	Yes, first 30 seconds used to redraw map and settings.
Should graphics such as the NAIs and phaselines be click/drag active during execution phase?	
Commander	Yes
Battlespace	Lock them down! When they move during execution, it is a pain.
Others	
Information Manager: *Need to put the toggles back for enemy units (by state).* *Need to add toggles for weapons (PAM, LAM, hits).* *Need different graphics for screen clarity.*	

Note: If a command group participant is omitted under a question, he recommended no change to the default setting.